DOCUMENT RESUME

ED 064 956	24	EM 010 206

TITLE	EIN Software Catalog.
INSTITUTION	Interuniversity Communications Council (EDUCOM), Boston, Mass.
SPONS AGENCY	National Science Foundation, Washington, D.C.; Office of Education (DHEW), Washington, D.C.
BUREAU NO	BR-8-0697/Attachment No. 1
PUB DATE	[72]
GRANT	OEG-0-8-080697-4468(007)
NOTE	623p.; See also EM 010 207, EM 010 208, EM 010 209, EM 010 210, EM 010 211, EM 010 212
EDRS PRICE	MF-$0.65 HC-$23.03
DESCRIPTORS	Behavioral Sciences; Biological Sciences; *Catalogs; Computer Oriented Programs; *Computer Programs; Earth Science; Educational Administration; Engineering; *Information Networks; Information Science; *Interinstitutional Cooperation; Library Science; Mathematics; Operations Research; Physical Sciences; *Program Descriptions; Statistics; Technology; Tests
IDENTIFIERS	*Educational Information Network; EIN

ABSTRACT

The EIN (Educational Information Network) is a non-profit operation which coordinates the sharing of educational computing resources. It is administered by EDUCOM and funded jointly by the U. S. Office of Education and the National Science Foundation. EIN maintains a group of contact personnel at member institutions to serve as a liaison between the institution and EIN. Through these persons items of software are offered for distribution. EIN also publishes a catalog of the software which is available through the network. The four-volume catalog contains an alphabetical listing of the participating EIN members that are represented in the catalog by program descriptions; descriptions of each computer facility listed and its general pricing algorithm; abstracts of available programs, subdivided into 13 areas of application; three indexes--by EIN number by descriptive title, and by keyword; and complete descriptions of programs, including user instructions, samples of input and output, and cost estimates. This first volume contains a description of EIN, facilities descriptions, abstracts, and indexes. Volumes two through four contain the complete program descriptions. (JY)

ED 064956

EIN SOFTWARE CATALOG

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEH REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

Prepared under a
Joint Grant from the
U.S. Office of Education
and the
National Science Foundation
Grant No. OEG-0-8-080697-[illegible]68 (007)

EM 010 206

EDUCOM 100 CHARLES RIVER PLAZA BOSTON, MASS. 02114 617-227-1805

The program descriptions that are contained in this catalog are based on information furnished to EDUCOM by the computer centers associated with EIN. EDUCOM shall have no liability whatsoever, including, without limitation, liability for errors contained in the description or arising out of the inappropriateness of the program for the purpose for which it is selected, to the user institutions. EDUCOM would appreciate being advised of any errors contained in these program descriptions.

The work presented or reported herein was performed pursuant to a Grant from the U.S. Office of Education, Department of Health, Education, and Welfare. However, the opinions expressed herein do not necessarily reflect the position or policy of the U.S. Office of Education, and no official endorsement by the U.S. Office of Education should be inferred.

12/69

EIN SOFTWARE CATALOG

Contents

THE EDUCATIONAL INFORMATION NETWORK
    The Purpose of EIN
    Eligibility
    How EIN Works
    The Intended Benefits of EIN
HOW TO USE THE EIN SOFTWARE CATALOG
    How to Find a Program
    How to Use a Program
    How to Submit a Request for Service
    How to Use the Network
    How to Use a Remote- Access System
    How to Supply Services to the Network
    EIN Network Overhead Charge
    How to Make Complaints and Comments
    How to Submit Software to the Catalog
LIST OF EIN MEMBERS
SECTION I Facilities
SECTION II Abstracts
SECTION III Index
SECTION IV Entries

3

THE EDUCATIONAL INFORMATION NETWORK

The Purpose of EIN

The Educational Information Network (EIN) is a nonprofit operation concerned with the sharing of educational computing resources. It is administered by EDUCOM and funded jointly by USOE and NSF. EIN is a network operating on the following premises.

(1) Software is used at the installation where it is currently running.

(2) The network is to distribute information about the availability of this software.

(3) The network is to act as administrative middleman between the user and the resource. It will secure the commitment of the resource to the user and will handle the billing and accounting.

(4) The network is in no way to interpose itself into the technical transaction.

Eligibility

Membership is open to any member of EDUCOM. Nonmembers may use the network by working through an EIN member at the discretion of that member.

How EIN Works

EIN maintains a group of contact personnel at member institutions, nationally and internationally, called the EIN technical representatives. Each of these persons is the liaison between his institution and EIN.

Through these persons, items of software are offered for distribution through the network. These include programs, systems of programs, facilities, interactive systems, etc. In other words, any item of software that is of interest to the academic community can be included.

EIN publishes writeups of the software in this periodical, called the *EIN Software Catalog*. It is supplemented monthly, updated as needed, and distributed to all institutions which are members of EIN. It makes all programs listed with EIN available for inspection.

Persons wishing to use a program make the arrangements through the EIN technical representative at their institution.

The Intended Benefits of EIN

In General:	Computational facilities are expensive. Technical abilities are in short supply. Requirements for many types of facility exist at the same institution. All these problems are helped by resource sharing.
To the Resource:	The *EIN Software Catalog* provides a medium for the publication and recognition of software. It is like a scholarly journal in that respect. In addition, EIN will notify the disciplinary journals of the inclusion of specific items. Usage of a facility by the network can provide better utilization by adding billable hours.
To the Users:	EIN vastly increases the computational power available. The sum of the facilities listed in the Catalog is a larger resource than any single institution can afford.

HOW TO USE THE EIN SOFTWARE CATALOG

The *EIN Software Catalog* is made up of four main sections.

I. Facilities

—includes (1) an alphabetical listing of the participating EIN members that are represented in the Catalog by program descriptions; (2) descriptions of each computer facility listed and its general pricing algorithm. New material for this section is sent with the update packages, as needed.

II. Abstracts

—is subdivided into 13 areas of application to facilitate the selection of programs; within subsections, functional abstracts of programs available for use are ordered by EIN Number. The 13 areas of application are:

1. Administrative
2. Behavioral Science
   (education, psychology, sociology, etc.)
3. Computer Utility
   (languages, utility programs, systems, etc.)
4. Earth Sciences
   (geography, geology, mineralogy, etc.)
5. Engineering & Technology
6. Humanities
   (art, history, music, etc.)
7. Library & Information Sciences
8. Life Sciences
   (biology, medicine, zoology, etc.)
9. Mathematics
10. Operations Research
    (linear programming, simulations, etc.)
11. Physical Sciences
    (astronomy, chemistry, physics, etc.)
12. Statistics & Measurement
13. General Facility Access
    (entire facility accessible)

The Abstracts Section is supplemented monthly along with the Entries Section.

III. Index
—is made up of three indices: by (1) EIN Number, (2) descriptive title, and (3) key words. The programs covered in the Keyword Index are those listed in the EIN Number Index bearing the same date. The Index Section will be made current at regular intervals and sent as part of the update packages.

IV. Entries
—contains more-complete descriptions of programs available for use than are given in the Abstracts Section; includes user instructions, samples of input and output, and cost estimates; programs are ordered by EIN Number. The Entries Section is supplemented monthly along with the Abstracts Section.

How to Find a Program

There are basically two ways to use this Catalog to find a program that might serve your needs. If you can find in the Keyword Index words or phrases that describe the operation that you want, the associated EIN Numbers can then be used to locate the pertinent program writeups in the Entries Section.

Alternately, by browsing through the relevant areas of application in the Abstracts Section, you should be able to determine the EIN Numbers of programs that may be of use to you. The more complete description of the Catalog entry for a particular program can then be located by the EIN Number in the Entries Section.

How to Use a Program

Material in the Entries Section is designed to contain enough information and instruction so that, in most cases, a program may be used immediately. It is generally *not* designed to delineate *all* of the available options and minor details of a given program.

If a program or system is too extensive or complex to allow a concise description, other literature is cited so that with the aid of your EIN technical representative, you can obtain further documentation. On the basis of the catalog entry, or with additional documentation, you should be able to decide upon your interest in a program and to prepare the necessary inputs for its use.

How to Submit a Request for Service

The first submitted request for a program to be run at one of the participating facilities requires contact through your EIN technical representative. He will have the necessary forms and instructions on how to establish an account with the selected participant. In addition, he may be of help in determining whether valid input data have been prepared.

How to Use the Network

Prior to any use of the network, a *credit* sufficient to cover the intended use must first be established with EIN. This can be done by issuing Purchase Orders, by depositing cash, or by depositing a cash-equivalent quantity of computing power. Each user must then establish an *account* for each resource that he wishes to use. For this purpose the EIN Account Initiation Form is provided.

Once the account is initiated, you work directly with the resource institution. Work is ordered by submitting an EIN Job Run Form directly to the resource EIN technical representative. Remember, however, that the work authorized may not exceed the amount remaining in the account.

At the end of each month, you will receive an account statement, listing your institution's total credits with the Network, the charges against each of your accounts, and the amount remaining in that account. You will also receive a separate invoice for each user, which can be used for internal purposes.

How to Use a Remote-Access System

The user will establish an account with EIN by submitting an EIN Account Initiation Form for each resource that he wishes to use. Upon validation, the EIN office will notify the resource institution of the initiation of the account.

The resource institution will then issue to the user the system-access instructions, log-in procedures, user codes, and special information. Once the user has acquired the information necessary to the use of a special remote-access system, he may then make use of the resource without having to submit individual Job Run Forms.

The resource institution will be responsible for ensuring that each user does not exceed his allowable usage. The resource will invoice the EIN office, which, in turn, will invoice the user.

How to Supply Services to the Network

You will be authorized to perform work for a given user upon receipt from EIN of a verified Account Initiation Form. You will receive individual job orders specified on the EIN Job Run Form.

When a job is completed, the work should be invoiced to EIN. You may use your standard accounting procedures. This invoice must be issued to EIN against the number appearing on the EIN Account Initiation Form.

Invoices should be directed to

Educational Information Network
EDUCOM
100 Charles River Plaza
Boston, Massachusetts 02114

EIN Network Overhead Charge

A charge of use of the EIN Network, amounting to 20% of the total resource charge, will be invoiced to the user.

How to Make Complaints and Comments

All complaints and comments (we hope that they will not all be complaints) should be channeled through your local EIN technical representative. He will be able to relay your remarks both to the EIN office and to the appropriate participant. He is also specially equipped to troubleshoot whenever necessary.

How to Get Additional Information

Your first source of additional information is your local EIN technical representative. He will be able to obtain information either from other participants or from the EIN office in Boston.

How to Submit Software to the Catalog

The network is prepared to consider any item of software that could be of use to the academic community. Both academic and administrative resources are eligible.

Software will be operated at the institution which submits it. The Catalog, therefore, does not list items which are intended only for export to the user.

Three categories of resources are being distributed through the network at this time: (1) programs; (2) systems of programs, including remote-access systems, and (3) entire facilities, accessible either electronically or otherwise.

Persons interested in submitting software in one of these categories should consult the *EIN Documentation Standards Handbook* and contact their EIN technical representative.

LIST OF EIN MEMBERS	TECHNICAL REPRESENTATIVE
Alabama, University of Tuscaloosa, Alabama	Mr. Roger B. Wicks
Alberta, University of Edmonton, Alberta	Dr. Dale H. Bent
Amherst College Amherst, Massachusetts	Mr. Paul J. Plourde
Arizona, The University of Tucson, Arizona	Dr. Roger van Norton
Boston University Boston, Massachusetts	Mr. John E. Alman
British Columbia, The University of Vancouver, British Columbia	Program Librarian
Brooklyn, Polytechnic Institute of Brooklyn, New York	Mr. Stanley Habib
Bucknell University Lewisburg, Pennsylvania	Mr. E. F. Staiano
Calgary, The University of Calgary, Alberta	Dr. Dennis Ojakangas
California, University of	
Berkeley, California	Mr. Kenneth Hebert
Davis, California	Mr. G. J. Kurowski
Irvine, California	Mr. Julian Feldman
Los Angeles, California	Mr. William B. Kehl
Riverside, California	Mr. Morris J. Garber
San Diego, California	Mr. Kenneth L. Bowles
San Francisco, California	Mr. John A. Starkweather
Santa Barbara, California	Mr. Charles R. Loepkey
Santa Cruz, California	Mr. Harry D. Huskey
Carnegie-Mellon University Pittsburgh, Pennsylvania	Mr. Harry Rowell
Catholic University of America, The Washington, D.C.	Dr. Andrew G. Favret
Cleveland State University, The Cleveland, Ohio	Mr. C. William Marcy III
Dalhousie University Halifax, Nova Scotia	Mr. John Howard Oxley

continued

LIST OF EIN MEMBERS	TECHNICAL REPRESENTATIVE
Dartmouth College Hanover, New Hampshire	Mr. A. Kent Morton
Dayton, University of Dayton, Ohio	Mr. Leo J. Chico
Drexel Institute of Technology Philadelphia, Pennsylvania	Prof. James B. Maginnis
Duke University Durham, North Carolina	Mr. Hamilton Hoyler
Educational Testing Service Princeton, New Jersey	Mr. Ernest Anastasio
Evergreen State College, The Olympia, Washington	Dr. Robert Barringer
Florida, University of Gainesville, Florida	Dr. R. G. Selfridge
Florida State University, The Tallahassee, Florida	Mr. Ray Soller
Georgia, The University of Athens, Georgia	Dr. James L. Carmon
Georgia Institute of Technology Atlanta, Georgia	Mr. A. P. Jensen
Harvard University Cambridge, Massachusetts	Dr. Norman Zachary
Hawaii, University of Honolulu, Hawaii	Miss Jean Foytik
Indiana University Indianapolis, Indiana	Dr. David A. Neal
Iowa, The University of Iowa City, Iowa	Mr. Phillip Dylhoff
Iowa State University of Science and Technology Ames, Iowa	Dr. Clair G. Maple
Kansas, University of Lawrence, Kansas	Mr. Paul J. Wolfe

continued

LIST OF EIN MEMBERS	TECHNICAL REPRESENTATIVE
Lehigh University Bethlehem, Pennsylvania	Dr. John E. Walker
Liege Universite de Liege, Belgium	Dr. A. Danthine
Maine, University of Orono, Maine	Mr. Jeremy Johnson
Marquette University Milwaukee, Wisconsin	Mr. Robert Miller
Maryland, University of College Park, Maryland	Dr. John P. Menard
McGill University Montreal, Quebec	Prof. W. D. Thorpe
Miami, University of Coral Gables, Florida	Mr. Bruce M. Weber
Michigan State University East Lansing, Michigan	Dr. Harry Eick
Michigan, Western, University Kalamazoo, Michigan	Prof. Jack Meagher
Middle Atlantic Educational & Research Center Lancaster, Pennsylvania	Dr. Paul W. Ross
Minnesota, University of Minneapolis, Minnesota	Dr. William Craig
Missouri, The University of Rolla, Missouri	Prof. Ralph E. Lee
Nevada, University of Reno, Nevada	Dr. Dan Oppelman
New Hampshire, The University of Durham, New Hampshire	Mr. Richard Burrows
New Mexico, The University of Alburquerque, New Mexico	Dr. Stoughton Bell
New York Institute of Technology New York, New York	Dr. Alexander Schure
New York University New York, New York	Prof. Max Goldstein

continued

LIST OF EIN MEMBERS	TECHNICAL REPRESENTATIVE
New York, The City University of New York, New York	Dean Robin E. Spock
New York, State University of Albany, New York	Mr. Harold Wakefield
New York, State University of Albany, New York	Mr. Robert J. Robinson
Northeastern University Boston, Massachusetts	Prof. Thomas E. Hulbert
Northwestern University Evanston, Illinois	Mrs. Lorraine Borman
Notre Dame, University of Notre Dame, Indiana	Mrs. Elizabeth Hutcheson
Oberlin College Oberlin, Ohio	Mr. Eric A. Sosman
Ohio College Library Center, The Columbus, Ohio	Dr. Frederick G. Kilgour
Oklahoma Medical Center, The University of Oklahoma City, Oklahoma	Dr. Arthur Nunnery
Pennsylvania, University of Philadelphia, Pennsylvania	Dr. Charles W. McClintock
Pennsylvania State University, The University Park, Pennsylvania	Dr. Daniel Bernitt
Pittsburgh, University of Pittsburgh, Pennsylvania	Mr. John Nold
Principia College Elsah, Illinois	Mr. William H. Miller
Quebec, Universite de Quebec, Quebec	Mr. Jean Paul Pelchat
Rochester Institute of Technology Rochester, New York	Mr. Ronald E. Stappenbeck

continued

LIST OF EIN MEMBERS	TECHNICAL REPRESENTATIVE
Southern Methodist University Dallas, Texas	Mr. James Stephens
Temple University Philadelphia, Pennsylvania	Mr. David Remis
Texas Tech University Lubbock, Texas	Mr. Thomas F. Lee
Vanderbilt University Nashville, Tennessee	Mr. Parker L. Coddington
Virginia, University of Charlottesville, Virginia	Dr. Alan P. Batson
Virginia Polytechnic Institute Blacksburgh, Virginia	Dr. Robert Heterick
Washington University St. Louis, Missouri	Dr. Jon C. Strauss
Washington State University Pullman, Washington	Dr. Mads Ledet
Wayne State University Detroit, Michigan	Dr. Charles H. Briggs
West Virginia University Morgantown, West Virginia	Dr. Wayne Muth
Wilkes College Wilkes-Barre, Pennsylvania	Dr. David Williams

FACILITIES

PARTICIPATING COMPUTER FACILITIES

California, University of
Santa Barbara, California

Catholic University of America, The
Washington, D.C.

Dalhousie University
Halifax, Nova Scotia

Dartmouth College
Hanover, New Hampshire

Educational Testing Service
Princeton, New Jersey

Florida State University, The
Tallahassee, Florida

Georgia, University of
Athens, Georgia

Indiana University—Purdue University at Indianapolis
Indianapolis, Indiana

Iowa, University of
Iowa City, Iowa

Iowa State University of Science and Technology
Ames, Iowa

Marquette University
Milwaukee, Wisconsin

Maryland, University of
College Park, Maryland

Michigan State University, The
East Lansing, Michigan

Middle Atlantic Educational and Research Center
Lancaster, Pennsylvania

Minnesota, University of
Minneapolis, Minnesota

New York, The City University of
New York, New York

Northwestern University
Evanston, Illinois

Notre Dame, University of
Notre Dame, Indiana

Pennsylvania, University of
Philadelphia, Pennsylvania

Pennsylvania State University, The
University Park, Pennsylvania

Pittsburgh, University of
Pittsburgh, Pennsylvania

Virginia, University of
Charlottesville, Virginia

Washington University
St. Louis, Missouri

8/71

NAME OF FACILITY	University of California, Santa Barbara Computer Center
NAME AND ADDRESS OF EIN TECHNICAL REPRESENTATIVE	Mr. Charles R. Loepkey, Assistant Director, Computer Center, University of California, Santa Barbara, California 93106 Tel.: (805) 961-2261
PRINCIPAL EQUIPMENT	IBM 360/75 IBM 360/20

STANDARD SYSTEM USE

The operating system in use is OS/MVT with HASP. Information on system usage can be found in the references listed under the specific software entry.

PRICING

Charges for use of the On-Line System are the sum of the following.

Description	*Price*
Connect time	$4.00/hr.
Central processor unit (CPU)	0.0001793/CRU[α]
Core units	0.1776/K-byte-hr.
Disk storage	0.10/K-byte-month

[α] A CRU is defined as CPU sec. X a program weighting factor which is dependent on the characteristics of the program.

The cost per hour varies from user to user but the average for all users is $8.40. Student usage is nearly always on the low side of the mean cost/hr.

NAME OF FACILITY	The Catholic University of America Computer Center
NAME AND ADDRESS OF EIN TECHNICAL REPRESENTATIVE	Dr. Andrew G. Favret, Dir., Computer Center, The Catholic University of America, Washington, D.C. 20017 Tel.: (202) 529-6000 ext. 661
PRINCIPAL EQUIPMENT	PDP—10

STANDARD SYSTEM USE

The PDP-10 is a timesharing system operating under the standard timesharing monitor. Details on use of the PDP—10 can be found in the reference listed below.

PRICE

Current charges for timesharing use are:

Central Processor Time	$150./hr.
plus	
Kilo-core Seconds	$ 20./hr. before 5 p.m. $ 10./hr. after 5 p.m.
plus	
Connect Time	$ 2.40/hr.

REFERENCES

PDP-10 User's Guide (Washington, D.C.: Catholic Univ. Comp. Ctr., March 1971). Available from Catholic University on request.

NAME OF FACILITY	The Dalhousie University Computing Center
NAME AND ADDRESS OF EIN TECHNICAL REPRESENTATIVE	John Howard Oxley, Dept. of Communications Services, I.W. Killam Memorial Library, Dalhousie University, Halifax, Nova Scotia, Canada
PRINCIPAL EQUIPMENT	CDC 6400

STANDARD SYSTEM USE

The operating System used is the standard SCOPE system. Control cards for the program run will be provided at run time.

PRICING

The present rate for computer time supplied to external educational users is \$150./processing hr (Canadian currency), where processing time is derived by the algorithm:

$$3(CP + \frac{PP}{2}) \text{ in seconds,} \qquad \begin{array}{l} CP = \text{CPU time} \\ PP = \text{peripheral processor time} \end{array}$$

and charges are computed as:

$$\frac{3(CP + \frac{PP}{2})}{3600} \times \$150.$$

The minimum charge is \$5.00 (Canadian), which includes postage and handling. Extra charges for special transmission modes, special forms or other custom work are available on request before job submission.

Non-educational external users' rate available on request.

NAME OF FACILITY	Kiewit Computation Center Dartmouth College
NAME AND ADDRESS OF EIN TECHNICAL REPRESENTATIVE	Mr. A. Kent Morton, Kiewit Computation Center, Dartmouth College, Hanover, New Hampshire 03755 Tel.: (603) 646-2864
PRINCIPAL EQUIPMENT	Dual GE-635

STANDARD SYSTEM USE

The GE-635 normally operates in a time-sharing mode, supporting such low-speed terminals as Teletype models 33, 35, and 37, Friden 7100, and IBM 2741. Other teletype-compatible and IBM-compatible devices will also be accepted by the system.

Dartmouth Time-Sharing (DTSS) is available from 8:00 a.m. to 1:00 a.m. on Monday, Tuesday, Thursday, Friday, and Saturday; from 8:00 a.m. to 8:00 p.m. on Wednesday; and from 8:00 p.m. to 1:00 a.m. on Sunday.

To be accepted by the system, a user must supply a valid user number which is provided when a new account is approved. Log-on procedures and other necessary information will be found in the reference below.

A potential user should first file an Account Initiation Form with the EIN Project Office. EIN will subsequently notify the Technical Representative at Dartmouth, who will contact the potential user and arrange for validating a passworded user number.

PRICING

	Educational Users	*Commercial Users*
Terminal connect time		
110 bps	$3.50/hr.	$11.00/hr.
150 bps	$4.35/hr.	$12.75/hr.
Central Processor	$0.11/sec.	$ 0.35/sec.
Storage	$2.95/1000 words	$ 3.00/1000 words

REFERENCES

Morton, A. Kent, *Concise Handbook to the Dartmouth Time-Sharing System*, (Kiewit Comput. Center, Dartmouth College, Hanover, N.H., June,1970). Available from the EIN Office at the cost of reproduction and mailing.

NAME OF FACILITY	Office of Data Analysis Research Educational Testing Service
NAME AND ADDRESS OF EIN TECHNICAL REPRESENTATIVE	Mr. Ernest Anastasio, Assistant Director, Office of Data Analysis Research, Educational Testing Service, Rosedale Road, Princeton, N.J. 08540 Tel.: (609) 921-9000 ext. 2552
PRINCIPAL EQUIPMENT	IBM System 360/65

STANDARD SYSTEM USE

The operating system in use is an OS/MVT with HASP.

PRICING

In a multi-programming environment, the elapsed time for a job is affected by other jobs that may be resident in the system. Consequently, the elapsed wall-clock time is not an equitable measure of usage. Recognizing the inadequacies of billing on wall-clock time, the new job billing algorithm is designed around a theoretical elapsed time called Computed Elapsed Time (CET). CET is an estimate of the elapsed job time through knowledge of the number of I/O-device accesses and the central processing unit (CPU) time for the job step.[1]

In brief, the cost for a job is computed according to the following equation.

$$C = D_p \times T_{cet} + T_{wct} \times \sum_i D_i + H + F$$

where:

C = Cost

D_p = Design partition hourly rate$^\alpha$

T_{cet} = Computed Elapsed Time

T_{wct} = Wall Clock Time

D_i = Hourly rate for the i-th *dedicated* device

H = HASP standard charge for I/O

F = Fixed charges (not presently used, but allows for flexibility)

$^\alpha$Design partition is the portion of the machine used for a job and can be determined from the table on the following page.

continued

CET Hourly Rates by Design Partition

	Tapes			
Core	*0*	*1-4*	*5-7*	*8 and up*
<20	$ 52.	$ 73.	$ 87.	$ 116.
21–40	106.	131.	145.	174.
41–60	158.	174.	196.	225.
61–80	210.	225.	247.	276.
81–100	264.	283.	305.	334.
101–120	316.	334.	355.	384.
121–140	370.	392.	413.	442.
141–160	422.	442.	464.	493.
161–180	476.	500.	515.	544.
181–200	528.	551.	566.	595.
201–220	580.	595.	616.	645.
221–240	632.	653.	674.	703.
241–260	687.	711.	732.	761.
261–300	792.	812.	834.	863.
301–350	924.	943.	964.	993.
351–400	1056.	1080.	1095.	1124.
401–500	1318.	1334.	1356.	1385.
>500	1450.	1450.	1450.	1450.

At present the rate for the full usage of the ETS computational facilities is $1,450/CET–hr.

Further information on the full pricing algorithm and details of an approximate algorithm can be obtained from the reference listed below. Interested persons should contact the EIN technical representative at ETS.

In addition to the basic computer charges, there will be a fee of $5.00 for postage and handling.

REFERENCE

1. Webb, J.H., *Systems and Programming Report SPG.TR.6: Machine Accounting and the Job Billing Algorithm* (Princeton, N.J.: Educational Testing Service, July 1970).

NAME OF FACILITY	The Florida State University Computing Center
NAME AND ADDRESS OF EIN TECHNICAL REPRESENTATIVE	Mr. Ray Soller, Computing Center, The Florida State University, Tallahassee, Florida 32306 Tel.: (904) 599-4770
PRINCIPAL EQUIPMENT	CDC 6400

STANDARD SYSTEM USE

The operating system used is the standard Scope system. Control cards (as dictated by the system) will, in general, be provided at run time.

PRICING

The algorithm used for computing charges is based on the number of computational units (hours) used for each device.

The number of computational units of computer time is obtained by adding the amounts of central-processor and peripheral-processor time used. The current rate for computer time is \$150./CU if the memory used is $\leq 70{,}000_8$, or \$225./CU if the memory used is $> 70{,}000_8$. (See note below.) Usage of I/O devices is charged at a rate of \$50./CU.

In addition, there will be a fee for consulting and handling (at a rate of \$7.50/h), with a minimum of one hour of the same being charged for each run.

Note: These rates do *not* apply when using the transportation package, for which the rate is set at \$400./CU.

NAME OF FACILITY	University of Georgia Computer Center
NAME AND ADDRESS OF EIN TECHNICAL REPRESENTATIVE	Dr. James L. Carmon, Director, Computer Center, University of Georgia, Athens, Ga. 30601 Tel.: (404) 542-3106
PRINCIPAL EQUIPMENT	IBM 360/65, IBM 7094, CDC 6400

STANDARD SYSTEM USE

The IBM 360/65 operates under OS 360, MVT, with HASP. The IBM 7094 is run under the IBSYS operating system. The CDC 6400 uses the standard Scope operating system.

PRICING

Pricing policies for the specific information search services from the University of Georgia announced in this catalog are available upon request.

NAME OF FACILITY	Indiana University—Purdue University at Indianapolis Research Computation Center
NAME AND ADDRESS OF EIN TECHNICAL REPRESENTATIVE	Dr. David A. Neal, Director, Research Computation Center, 1100 West Michigan St., Indianapolis, Ind. 46202 Tel.: (317) 264-7951
PRINCIPAL EQUIPMENT	IBM 7040 (batch mode) 32K words memory PDP-8-I (time-sharing mode) 20K words memory

STANDARD SYSTEM USE

The IBM 7040 and the PDP-8-I at IU Med. Center, Indianapolis, are offered to EIN users, with RCC personnel available for consultation in utilizing the Library, keypunching, etc. Executive system of the 7040 is IBSYS. Control Cards are described in Refs. 1 and 2. Jobs may be submitted to the 7040 from 8:30 to 5:30, Monday through Friday. Output is usually available within 24 hours.

The PDP-8-I (TSS-8) is available 24 hours a day, except long weekends. It can be accessed by any ASCII terminal operating at 110 baud, full duplex. Access instructions can be obtained from the EIN Technical Representative. User instructions can be found in Ref. 3.

The 7040 can be programmed in FORTRAN IV, COBOL, MAP and COMIT. Peripherals available for use by the 7040 include two disk modules with 56 million characters, four IBM 729 VI magnetic tape drives, an IBM 1403 printer, an IBM 1402 card reader/punch and an 11' CalComp plotter.

The PDP-8-I can be programmed in FOCAL, BASIC and PAL-D. Peripherals available include a 256K disk module, 2 DEC tape drives and a paper tape punch/reader.

PRICING

Prices are for educational use only.

Description	*Cost*
IBM 7040	
Computer Time	$90.00/hr.

continued

Description	*Cost*
Programming	
Basic	$7.50/hr.
General	$10.00/hr.
Complex	$15.00/hr.
PDP-8-I	
CPU Time	$3.75/min.
Device Hours	$3.00/hr.
Hybrid System	$30.00/hr.

REFERENCES

1. *Programmer's Guide*, CG28-6318-8 (White Plains, N.Y.: IBM).

2. *Programmer's Guide Supplement*, RCC #20.0 (Indianapolis: IUPUI RCC).

3. *Introduction to Programming* (Maynard, Mass: Digital Equip. Corp.).

NAME OF FACILITY	University Computer Center The University of Iowa
NAME AND ADDRESS OF EIN TECHNICAL REPRESENTATIVE	Mr. Phillip A. Dylhoff, Head, Applications Programming, University Computer Center, The University of Iowa, Iowa City, Iowa 52240 Tel.: (319) 353-3832
PRINCIPAL EQUIPMENT	IBM System 360/65

STANDARD SYSTEM USE

The operating system used is the OS/MVT with HASP II. Control cards for the running of programs will, in general, be provided at run time.

PRICING

Processor Charges

Device	*University rate*	*Commercial rate*
Central processor unit (CPU)	$125./hr.	$260./hr.
Core storage		
Main	$0.25/megabyte-sec.	$0.37/hr.
Slow*	$0.06/megabyte-sec.	$0.10/megabyte-sec.
Disk access	$0.0003/access	$0.0005/access
Tape access	$0.005/device-sec.	$0.016/device-sec.

I/O Charges

Device	*University rate*	*Commercial rate*
Terminal access	$1.75/connect-hr.	$3.00/connect-hr.
Card input		
Local	$0.0005/card	$0.0008/card
Remote	$0.0002/card	$0.0005/card
Card output	$0.0015/card	$0.0018/card

*: not available for batch jobs

continued

Device	*University rate*	*Commercial rate*
Printing		
Local	$0.0003/line	$0.0007/line
Remote	$0.0002/line	$0.0003/line
Plotting	$14./hr.	$20./hr.
Paper tape access	$0.36/device-sec.	$0.05/device-sec.
Supplies	$0.0035/page	$0.0035/page

NAME OF FACILITY	Iowa State University Computation Center
NAME AND ADDRESS OF EIN TECHNICAL REPRESENTATIVE	Dr. Clair G. Maple, Director, Computation Center, Iowa State University, Ames, Ia. 50010 Tel.: (515) 294-3402
PRINCIPAL EQUIPMENT	IBM System 360/65

STANDARD SYSTEM USE

The operating system used is the OS/MVT (not under HASP). Control cards for the running of programs will, in general, be provided at run time.

PRICING

The rate schedule for usage of computational facilities is very simple. The rate for expended CPU time is $375./h; in real time, however, the rates for core storage are $125./h for main core (per 256K bytes) and $17.50/h for large core storage (per 256K bytes).

All charges incurred will be subject to a minimum of $5.00/run. This includes postage, handling, and consulting fees. For runs that require greater amounts of computer time (for which computer charges alone exceed $5.00), there will be no additional charges. Thus, for larger jobs, the total charge will be comprised only of the specific computer usage costs.

NAME OF FACILITY	Marquette University Computing Center
NAME AND ADDRESS OF EIN TECHNICAL REPRESENTATIVE	Mr. Robert Miller, Acting Director, Computing Center, Marquette University, 1515 W. Wisconsin Avenue, Milwaukee, Wisc. 53233 Tel.: (414) 224-7700
PRINCIPAL EQUIPMENT	IBM 7040

STANDARD SYSTEM USE

The executive system used is the IBSYS system. Control cards are described in the *User's Manual: MUCC 00*, and the *Programmer's Guide: MUCC 10.*

PRICING

Charges for computational services for small colleges may be waived. The Computing Center personnel are available for consultation, programming assistance, and keypunching without charge on a limited basis; extensive utilization of these services can be obtained on a contractual basis.

Service	*Hourly Rate*
Computation	$240.00
Consultation	$ 10.00
Programming	$ 15.00
Keypunching	$ 2.50

NAME OF FACILITY	The University of Maryland Computer Science Center
NAME AND ADDRESS OF EIN TECHNICAL REPRESENTATIVE	Dr. John P. Menard, Associate Director, Computer Science Center, University of Maryland, College Park, Md. 20742 Tel.: (301) 454-0100
PRINCIPAL EQUIPMENT	UNIVAC 1108, IBM 7094

STANDARD SYSTEM USE

The UNIVAC 1108 is run under the EXEC 8 operating system while the IBM 7094 executive system is IBSYS. The 1108 has both remote and local batch capacity as well as demand processing capability. The 7094 is a batch-oriented system.

PRICING

The UNIVAC 1108 charge is $720./hour of CPU usage with a minimum of 2 seconds per run. The 7094 charge of $222./hour is based on the total time the job was in the system. Excessive consumption of cards and paper during output may result in an additional charge.

In addition to the computer usage charges, a $15. postage and handling fee will be charged for each job that originates at another institution.

NAME OF FACILITY	The Michigan State University Computer Laboratory
NAME AND ADDRESS OF EIN TECHNICAL REPRESENTATIVE	Dr. Harry Eick, Professor of Chemistry, Michigan State University, East Lansing, Michigan 48823 Tel.: (517) 355-5047
PRINCIPAL EQUIPMENT	CDC 6500 CDC 3600

STANDARD SYSTEM USE

The operating system used is the standard Scope system. System Control Cards will be provided by The Michigan State University Computer Laboratory Applications Programming group.

PRICING

Computer Rate Structure*

3600 System

Central processor	$245/hour
Card read	$0.50/1000 cards
Card punch	$0.75/1000 cards
Plot	$14.40/plotter hour
Print	$0.20/1000 lines

6500 System

Central process	$175/hour
Central memory	$0.002/octal word—CPU hour
Peripheral processor	$25/hour
Connect time	$4/hour
File storage	$0.03/PRU—month
Card read	$0.50/1000 cards
Card punch	$0.75/1000 cards
Print	$0.20/1000 lines

Extra paper and card charges will be assessed for all classes of service in the following manner,

Card punch	$1.10/1000 cards for all output greater than 250 cards per problem number per month

*Billings made to a non-University account number will be assessed a surcharge of 10%.

continued

Print	$0.005/page for all output greater than 250 pages per problem number per month
Plot	$0.01/inch for all output

Supplies and Miscellaneous Services

Keypunch	$4.50/hour
Verify	$4.50/hour
Tape storage	$0.50/month
1200 ft. magnetic tape	$13.00/each
2400 ft. magnetic tape	$16.75/each

Applications Programming

Decks submitted through EIN to the Michigan State University for processing will be handled by the Applications Programming group. A fee of $10.00 per hour is charged for this service.

NAME OF FACILITY	Middle Atlantic Educational and Research Center
NAME AND ADDRESS OF EIN TECHNICAL REPRESENTATIVE	Mr. Paul W. Ross, Manager of Services, MERC, P. O. Box 1372, Lancaster, Pa. 17604 Tel.: (717) 393-0132 or 393-5021
PRINCIPAL EQUIPMENT	RCA Spectra 70/46G

STANDARD SYSTEM USE

The RCA Spectra 70/46G normally operates in a time-sharing mode under TSOS, supporting such low-speed terminals as Teletype Model 33 or 35.

MERC facilities are available 8:00 a.m. to midnight Monday through Friday and 10:00 a.m. to 6:00 p.m. on Saturday, except for scheduled maintenance time 8:00 a.m. to 10:00 a.m. on Tuesday and Friday. Other hours may be specially arranged.

MERC is using all language capabilities offered by RCA: Assembly, COBOL, FORTRAN IV, BASIC, Interactive FORTRAN (IFOR), and COBOL Syntax Checker. Interested persons should consult the contact person.

PRICING

Description	*Cost*
CPU Time	$0.05/CRU[α]
Connect Time	
110 baud to 150 baud	$ 9.00/hr.
300 baud	$12.00/hr.
1200 baud	$15.00/hr.
2000 baud	$20.00/hr.
Peripheral Time	
Card reading	$0.20/1000 cards
Card punching	$2.00/1000 cards
Line printing	$0.75/1000 lines
Private volume usage (tape & disk)	$0.10/minute

[α] A CRU is defined as CPU seconds times a program weighting factor which is dependent upon the characteristics of the program.

continued

Description	*Cost*
Storage Charges On Line Storage (590 Disks)	
First 25 pages*	No charge
Next 75 pages	$1.00/page
Next 400 pages	$0.70/page
Next 500 pages	$0.50/page
Above 1000 pages	$0.25/page

*A page is 2048 characters (8 bit + parity) of storage

Off Line Storage (Magnetic tape)	
MERC supplied tapes	$2.00/month
Customer supplied tape	$1.00/month

Guaranteed Minimum Discount Policy

If the user is willing to guarantee MERC a monthly minimum billing, MERC will give the user a discount on computer charges (i.e., CPU Time, Connect Time, and Peripheral Time) in the amount of (minimum guaranteed by customer/100)% up to a maximum of 30%.

Educational users fall under the following minimum discount policy. There is no maximum to the amount of discount available and the discount is calculated as:

$$(\sqrt[3]{\text{minimum guaranteed/month}})\%$$

Educational Discount

Because of MERC's dedication to the educational community, a very liberal discount of 50% is allowed to educational institutions. This discount is applied to the total billing after any guaranteed minimum discounts are applied. However, educational institutions electing the guaranteed minimum discount will be expected to make the minimum amount after all discounts are applied.

Costs of manuals, consulting, supplies, and other support costs are not covered by the above discount policies.

NAME OF FACILITY	University of Minnesota University Computer Center
NAME AND ADDRESS OF EIN TECHNICAL REPRESENTATIVE	Mr. William J. Craig, Assistant Director, CURA, 311 Walker Library, University of Minnesota, Minneapolis, Minn. 55455 Tel.: (612) 373-7833
PRINCIPAL EQUIPMENT	CDC 6600

STANDARD SYSTEM USE

The operating system in use is the SCOPE system. Description of the necessary control cards and additional information of use of the University of Minnesota facilities can be found in the reference cited below.

PRICING

Description	*Price*
Central processor time	
Educational usage	$12./min.
Other usage	20./min.
Printed output	.02/page

In addition, charges for postage and handling will be made commensurate with the costs incurred.

REFERENCE

Computer User's Manual (Minneapolis: Univ. of Minn. Comp. Center, Jan. 1971).

NAME OF FACILITY	The City College of The City University of New York Computation Center
NAME AND ADDRESS OF EIN TECHNICAL REPRESENTATIVE	Dean Robin E. Spock, The City University of New York, 535 East 80 Street, New York N.Y. 10021 Tel.: (212) 360-2187 George W. Elder, Computation Center, The City College of The City University of New York, 139 Street & Convent Avenue, New York, N.Y. 10031 Tel.: (212) 621-2374
PRINCIPAL EQUIPMENT	IBM System 360/50 I, IBM 7040

STANDARD SYSTEM USE

The IBM 360/50 I is run under IBM's O.S./360 and the IBM 7040 under the IDSYS.

PRICING

The cost of using the IBM 360/50 I is computed according to the following formula.

$$\text{COST} = [\text{ET} + (\frac{\text{CR}}{15} + \frac{\text{LP}}{20} + \frac{\text{CP}}{5})/100] \times \$2.50 \text{ (per minute)},$$

where ET is execution time (including CPU time and voluntary wait time), CR is cards read, LP is lines printed, and CP is cards punched.

Use of the IBM 7040 is at the rate of $150/h.

NAME OF FACILITY	Vogelback Computing Center Northwestern University
NAME AND ADDRESS OF EIN TECHNICAL REPRESENTATIVE	Lorraine Borman, EIN Technical Representative, Vogelback Computing Center, Northwestern University, 2129 Sheridan Road, Evanston, Ill. 60201 Tel.: (312) 492-3682
PRINCIPAL EQUIPMENT	CDC 6400

STANDARD SYSTEM USE

The Scope 3.1 operating system is used. Two special cards are used to subdivide a job deck. An end-of-record (EOR) card contains a 7-8-9 multipunch in Col. 1. An end-of-information (EOI) card contains a 6-7-8-9 multipunch in Col. 1. Generally, the remainder of these two cards should remain unpunched.

A job consists of a deck of punched cards, beginning with a sequence card (supplied at Vogelback Computing Center) and ending with an EOI card.

Sequence card (supplied by center)
Job card ⎤
C_2 ⎥ control cards
⋮ ⎥
C_n ⎦
7-8-9 (EOR)
Data Deck 1
7-8-9 (EOR)
Data Deck 2
7-8-9 (EOR)
⋮
Data Deck i
6-7-8-9 (EOI)

PRICING

Charges for usage are based on the amount of central-processor and peripheral-processor time used and the amount of central memory occupied. The total computer charge is given by the formula

$$\text{charge} = R_i \times [\text{CP time} + (0.20 \times \text{PP time})],$$

continued

where:

R_1 TABLE

	Job Field Length		
	0–43K	43K–100K	>100K
Regular Job	$7.50/min.	$8.50/min.	$9.50/min.
Long Job[a]	$6.50/min.	$7.00/min.	$7.50/min.

The minimum computer charge for each job run at the central site is $1.00. The minimum charge for remote terminal jobs is $0.50. In addition to the computer charge, there is a $15.00 handling and postage fee for each job that originates at another institution.

---

[a] Long jobs are those printing over 200 pages of output or those using more than 10 minutes of central-processor plus peripheral-processor time with a core usage of over 70,000 words. Most EIN jobs will not be subject to this rate.

NAME OF FACILITY	The University of Notre Dame Computing Center
NAME AND ADDRESS OF EIN TECHNICAL REPRESENTATIVE	Mrs. Elizabeth Hutcheson, Computing Center, University of Notre Dame, Notre Dame, Ind. 46556 Tel.: (219) 283-7784
PRINCIPAL EQUIPMENT	UNIVAC 1107 with 65,536 36-bit words of core memory and 786,432 36-bit words of drum storage; 14 Uniservo IIA tape units.

STANDARD SYSTEM USE

The operating system used is the EXEC II system. Control cards will be provided at run time.

PRICING

Charges are based on total CPU time used during one calendar month and input-output usage.

Central Processor Charges

0 – 5 hrs.		$480.00/hr
5 – 10 hrs.	480.00 less 5% =	$456.00/hr.
10 – 15 hrs.	" " 10% =	$432.00/hr.
15 – 20 hrs.	" " 15% =	$408.00/hr.
20 – 25 hrs.	" " 20% =	$384.00/hr.
use in excess of 25 hrs.	" " 25% =	$360.00/hr.

The above rates for extended use of the computer do not apply to the total number of hours used. That is, the rate for the first five hours is $480.00/hr.; for the next five hours of use, $456.00/hr.; the next five hours, $432.00/hr.; etc.

Input/Output Charges

Card reader:	5 cards read for 1¢ (rounded down)
Card punch:	1 card punched for 1¢
Printer:	1 page (single-ply) paper printed for 5¢ (additional charges will be made for printing on multiple-part paper.)

In addition to the basic computer charge, there will be a fee of $10.00 for postage and handling.

NAME OF FACILITY University of Pennsylvania Computer Center

NAME AND ADDRESS OF EIN TECHNICAL REPRESENTATIVE
Dr. Charles W. McClintock, Director of Operations, The Computer Center, University of Pennsylvania, 3401 Market Street,
Philadelphia, Pennsylvania 19104
Tel.: (215) 594-5841

PRINCIPAL EQUIPMENT IBM System 360/75

STANDARD SYSTEM USE

The operating system is comprised of HASP II (Houston Automatic Spooling with Priority) and MFT II (Multiprogramming with a Fixed Number of Tasks), spooling input to disk prior to execution and output back to disk after execution, so as to maximize the flexibility, i.e., minimize the expended time, with regard to I/O. Jobs are partitioned into priority classes to allow greater efficiency in scheduling jobs for processing. HASP also has the capability of remote job entry.

In this hybrid operating system, a job has two priorities: an O.S. (dispatching) priority and a HASP (scheduling) priority. The O.S. priority depends on the partition in which the job is to be executed, with partition 3 (P_3) being the one with highest priority for *users*, having 110K bytes of storage. [Partition 4 (P_4) has 200K bytes.]

The overall HASP priority is determined by two things: (1) the HASP job class and (2) execution and I/O requirements. At present, there are seven of these job classes, summarized in the table below and listed in their normal order of priority (decreasing).

Class	*Usual Partition*	*Use*	*Time Limit*	*Line Limit*	*Punched-Card Limit*
W	3	All *express* WATFOR jobs with no punched output	10 sec	1000	0
X	3	All *express* jobs except WATFOR	1 min	6000	2000

continued

Class	*Usual Partition*	*Use*	*Time Limit*	*Line Limit*	*Punched-Card Limit*
Z	4	All WATFOR jobs to be run in main batch (no punched output)	10 sec	1000	0
C (premium)	4	Any O.S. job; short jobs processed first (with 2 min, 5000 lines, and 600 cards being default limits when no other limits are specified)	60 min	50 000	9999
A (base)	4	Any O.S. job (same default limits as class C)	60 min	5000	9999
E (standby)	4	Any O.S. job (same default limits as class C)	100 min	100 000	9999
F (lastrun)	4	Any O.S. job (same default limits as class C)	100 min	100 000	9999

Note: Classes E and F are run only when there are no jobs in any of the other classes.

PRICING

Pricing policies for the specific University of Pennsylvania programs entered in this Catalog are listed as part of each program writeup.

NAME OF FACILITY	The Pennsylvania State University Computation Center
NAME AND ADDRESS OF EIN TECHNICAL REPRESENTATIVE	Dr. Daniel L. Bernitt, The Pennsylvania State University Computation Center, 105 Computer Building, The Pennsylvania State University, University Park, Pennsylvania 16802 Tel.: (814) 865-9527
PRINCIPAL EQUIPMENT	IBM System 360/67

STANDARD SYSTEM USE

The System 360/67 is run under IBM's O.S./360 and a locally developed Remote Job Entry (RJE) system, supporting such low-speed terminals as the IBM 2741, IBM 1050, and DATEL THIRTY-21. More detailed information as to usage of the RJE system may be obtained from the manual "Remote Job Entry System—General Information and Summary Description," available from the EIN office at the cost of duplication and mailing.

PRICING

Whether a batch-processing job originates directly (through high-speed entry) or indirectly (via the RJE system), charges for computer time on the 360/67 are calculated at the rate of $0.11/sec.

When the RJE system is used, there are, in addition, two types of specific charges. First is the charge for usage of a typewriter terminal, based upon the actual connection time (time between "logging on" and "logging off"). The amount of time is calculated to the nearest minute and charged at a rate of $0.07/min.

The second charge is for the use of storage space, computed on the basis of the number of files (where a file consists of 500 eighty-column card images) belonging to a user and the number of days each such file is retained. In general the retention of files is charged at a rate of $0.12/file/day.

NAME OF FACILITY	University of Pittsburgh Computer Center
NAME AND ADDRESS OF EIN TECHNICAL REPRESENTATIVE	John Nold Assistant Director for Services Computer Center, University of Pittsburgh 800 Cathedral of Learning Pittsburgh, Pennsylvania 15213 Tel.: (412) 621-3500, ext. 7185
PRINCIPAL EQUIPMENT	IBM 7090 IBM System 360/50 (2)

STANDARD SYSTEM USE

The IBM 7090 and one of the 360/50 computers are utilized to provide batch-processing service. The 7090 uses a modified version of the monitor-controlled University of Michigan Executive System, while the model 50 is controlled by IBM's O.S./360. Time-sharing capability is available from the second 360/50, operating under the University of Pittsburgh Time-Sharing System (PTSS), a system which has the added feature of concurrent, background batch-processing (for larger programs).

Use of the Time-Sharing System requires "signing on," by typing $$ LOGON USERID, where the term *userid* refers to an assigned project number followed by an individual "man-number" code. Upon acceptance of the *logon*, the user may freely utilize PTSS. (For detailed usage, it will be necessary to obtain the reference given below.) When he is finished, he so indicates by typing $$ LOGOFF, at which time the necessary accounting information is updated.

To use the 7090 system or O.S./360, the approved project number and man-number code are punched onto an ID card which must precede all other cards in a submitted deck.

REFERENCES

Pitt Time-Sharing System for IBM System/360 (Univ. Pittsburgh Computer Center, Pittsburgh, Pa.; revised in Mar. 1968).

PRICING

Charges for batch-processing service are computed at fixed hourly rates ($250/h for the 7090 and $225/h for the 360/50). However, the total charges resulting from use of the PTSS are determined by the amount of usage of each of three system components:

continued

(a) CPU time, in tenths of a second;

(b) core storage, determined by the product of the number of used blocks (where a block of core is 8192 bytes) and the execution time to the nearest 1/300 sec;

(c) I/O device usage, adding the individual device time-billing factors, each of which is calculated to the nearest 1/300 sec.

The overall charges for the time-sharing service provided are determined, then, by summing the products of these three major billing factors and their respective billing rates. The detailed rates are not included in this description; however, a reasonably accurate estimate of cost may be acquired by calculating time-sharing charges at a rate of $247/h.

NAME OF FACILITY	The University of Virginia Computer-Science Center
NAME AND ADDRESS OF EIN TECHNICAL REPRESENTATIVE	Dr. Alan Batson, Director, Computer-Science Center, Gilmer Hall, The University of Virginia, Charlottesville, Va. 22903 Tel.: (703) 924-3731
PRINCIPAL EQUIPMENT	Burroughs B5500

STANDARD SYSTEM USE

The operating system used is a slightly modified form of the Burroughs Mark 10 multiprogrammed batch processing system. Control card information for normal runs is listed in the Center's *Users Manual* which is available from the Center. The primary language is ALGOL, with all the other regular languages also available.

PRICING

It is expected that the B5500 will be replaced with a newer, larger system in the near future. Therefore only short-term projects that cannot effectively be processed elsewhere will be accepted.

	Prime Shift	*Third Shift*
Processor time	$116./hr	$70./hr
I/O time	58./hr	35./hr
Consulting service	7./hr	
Plotter time	13./hr	

Nominal charges are made for disk storage, tape rental, keypunching, etc.

The above charges represent the cost to other educational institutions only.

NAME OF FACILITY	The Washington University Computing Facilities
NAME AND ADDRESS OF EIN TECHNICAL REPRESENTATIVE	Dr. Jon C. Strauss, Director, Inform. Processing Ctr., Sever Hall, Washington University, St. Louis, Mo. 63130 Tel.: (314) 863-0100 ext. 3102
PRINCIPAL EQUIPMENT	IBM System 360/50

STANDARD SYSTEM USE

The S/360 is run under IBM's OS/360 in a MFT-HASP environment, OS JCL will, in general, be provided at run time.

PRICING

Charges for S/360 usage are computed from the following formula:

S/360 Charge = CPU * PF * R1 + CI * R2 + CO * R3 + LO * R4

where:

CPU = The total CPU time spent in processing the job. This does not include wait time, scheduler time, or tape or disk mounting time.

PF = Partition factor. This factor is based upon the resources of the machine (primarily core used) which are devoted to the job. The values are:

Job Class	PF	Partition Size (K=1024 bytes)	High or Low Speed Core
A*	1.000	100K	high
B	.400	50K	high
C	.300	100K	low
E	1.250	150K	high
F	.375	150K	low
G	.450	200K	low
H	.525	250K	low

*Default class value.

R1 = The CPU rate. Currently set at $8.40/minute for contract research and $11.70/minute for industrial sponsored research.

continued

CI = Cards in. The total number of cards of input submitted with the job.

R2 = Rate for cards in. Currently set at $0.0002/card.

CO = Cards out. The total number of cards produced as punched output from the job.

R3 = Rate for cards punched. Currently set at $0.0006/card.

LO = Lines out. The total number of lines of output produced by the job.

R4 = Rate for lines of printed output. Currently set at $0.0002/line.

A fee will also be charged for handling and consultation at the rate of $10.00 per hour.

The total cost = S/360 charges + handling charges + postage

A
B
S
T
R
A
C
T
S

Administrative

DESCRIPTIVE TITLE	Information-Oriented Language—A Generalized Information and Retrieval System
CALLING NAME	INFOL
INSTALLATION NAME	Vogelback Computing Center, Northwestern University
AUTHOR(S) AND AFFILIATION(S)	Dr. Jacques Vallee, Control Data Corporation Robert Chalice Department of Industrial Engineering, Northwestern University (CDC 6400 version)
LANGUAGE	CDC FORTRAN IV
COMPUTER	CDC 6400
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Lorraine Borman, EIN Technical Representative, Vogelback Computing Center, Northwestern University, 2129 Sheridan Road, Evanston, Ill. 60201 Tel.: (312) 492-3682

FUNCTIONAL ABSTRACT

INFOL (Information-Oriented Language) is a generalized information storage and retrieval system that permits the user to extract selected pieces of information from a file, presented quickly and conveniently in their original form.

The INFOL language, a subset of English, was designed to be learned and used easily by those without a programming background or a knowledge of computers.

INFOL is a free-field language, thus eliminating the requirement for any part of the language to appear in specific columns of a punched card. Although punched cards are used, the concept of cards as unit records is generally avoided. The concept of an information field, implying limited word size, has been eliminated in INFOL; the user need not be concerned about how many characters to assign any piece of information. All information used as input is treated as a variable-length, but continuous, string of characters. The output may be a printed report or it may be a punched-card deck to be used as input for another computer application.

continued

INFOL includes several basic information-retrieval operations. It first allows the user to store (and later to modify) quantities of data. The system requires no input commands; files of information are established by the user with a single control word, followed by data descriptions and the data. INFOL also facilitates the selection and retrieval of information from the files according to specific requirements (retrieval criteria). The comprehensive examination (interrogation) of stored data, based on efficient retrieval criteria, is the heart of the INFOL system.

REFERENCES

INFOL for the CDC 6400 is based on INFOL as developed by the Control Data Corporation for the CDC 3600/3800 computer. Northwestern University gratefully acknowledges CDC's support in permitting it to utilize INFOL as the basis for the 6400 version and for permission to use the INFOL reference manual number 60170300 to produce its own manual, "INFOL for the CDC 6400—Information Storage and Retrieval System" (Nov. 1968). Support was also given by the U.S. Air Force Office of Scientific Research: "Man-Machine Interaction in Information Retrieval Systems."

000 0029

DESCRIPTIVE TITLE	Critical-Path Summary—Probability Statistic Based on Normal Curve
CALLING NAME	PERTC NUCC022
INSTALLATION NAME	Vogelback Computing Center Northwestern University
AUTHOR(S) AND AFFILIATION(S)	Vogelback Computing Center Northwestern University
LANGUAGE	CDC FORTRAN
COMPUTER	CDC 6400
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Lorraine Borman, EIN Technical Representative, Vogelback Computing Center, Northwestern University, 2129 Sheridan Road, Evanston, Ill. 60201 Tel.: (312) 492-3682

FUNCTIONAL ABSTRACT

PERTC produces a detailed scheduling network, final critical-path summary, and probability statistic for events that are normally distributed about some mean.

REFERENCES

Rogers, H.B., "Construction Project Pert Network," Northwestern Univ. Grad. School Business (Jan. 1964; unpublished). A reference copy may be obtained from the EIN office for the cost of duplicating and mailing.

000 0035

DESCRIPTIVE TITLE	Test Scorer and Statistical Analysis
CALLING NAME	TSSA NUCC072
INSTALLATION NAME	Vogelback Computing Center Northwestern University
AUTHOR(S) AND AFFILIATION(S)	Richard Wolf and Leopold Klopfer The University of Chicago K. Jones Harvard University Graduate School of Education A. Gasche, B. Wright, and C. Bradford The University of Chicago
LANGUAGE	CDC FORTRAN IV
COMPUTER	CDC 6400
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Lorraine Borman, EIN Technical Representative, Vogelback Computing Center, Northwestern University, 2129 Sheridan Road, Evanston, Ill. 60201 Tel.: (312) 492-3682

FUNCTIONAL ABSTRACT

TSSA scores multiple-response tests that have a single correct response for each item, computes test and item statistics, computes the tetrachoric interitem correlation matrix, and performs a factor analysis and varimax rotation. The point-biserial correlation of each item with a criterion score also may be obtained. The number of alternative responses may vary from item to item.

Multiple-scoring keys may be used with the same set of data cards. This feature makes TSSA applicable for scoring and analysis of the Kuder preference record and similarly constructed multikeyed instruments. In addition, multiple jobs may be run.

Computational results include the following.

Individual Scores, including raw scores and scores corrected for guessing

continued

Item-Response Information, including the proportion of subjects selecting the correct response for each item (i.e., difficulty) and the proportion of subjects selecting each response

Test Statistics,[1] including the mean, standard deviation, skewness, and kurtosis (and their standard errors computed from formulas outlined by Fisher); test reliability (computed by Kuder-Richardson formula 20); a validity coefficient; and a Pearson product-moment correlation

Item-Analysis Information,[2,3] including the proportion of subjects passing an item, the item standard deviation, point-biserial correlations, reliability index, and validity index (computed from formulas given by Guilford and Gullicksen)

Tetrachoric Interitem Correlations[4]

Factor Analysis of Interitem Correlation Matrix,[5] with varimax rotation and plot

Multiple-Scoring Keys

A special feature of TSSA is the use of Multiple-Scoring Keys. Such a feature has two uses. (1) It allows a single set of responses to be scored in more than one way; practical application of this is the use of the program with data from an instrument such as the Kuder preference record where several scores, based on analyses of the same set of items, are desired. (2) A more common situation is the use of the program to obtain subtest scores and a total test score from one instrument. An application of this might be to score all the responses on a reading test and to obtain scores for reading speed, reading vocabulary, and level of comprehension, as well as a total reading score.

Multiple Jobs

The limitation of the procedures in the analyses of subtests is that only the raw scores for each individual or each subject will be meaningful. Corrected scores will *not* have any meaning. Thus, if an analysis excludes any items initially read, no corrected score will be printed. The reason for this is that the program treats *all* items excluded from analysis as incorrect responses. Thus, the use of a formula for correcting scores for guessing is fallacious in this instance. If the user desires meaningful corrected scores for subtests, he will have to reproduce his Data Cards and run multiple jobs, selecting for reading on any one job only those items for which analysis is desired, i.e., the subtest.

continued

The use of the Multiple-Jobs feature for scoring subtests yields meaningful corrected scores, as already mentioned. However, the advantage of the Multiple-Scoring-Key feature of the program is that the data need be read from cards only once. Subsequent reading of the data is done from binary tape (tape B3), thus achieving savings of time and, hence, cost.

The Multiple-Jobs feature may also be used for scoring and analyzing severly entirely different tests, with only a single loading of the program.

REFERENCES

1. Fisher, R.A., *Statistical Methods for Research Workers* (Oliver & Boyd, Edinburgh, 1954), 12th ed., pp. 70-75.

2. Guilford, J.P., *Psychometric Methods* (McGraw-Hill Book Co., Inc., New York, 1954), pp. 373-464.

3. Gullicksen, H., *Theory of Mental Tests* (John Wiley & Sons, Inc., New York, 1950), pp. 363-396.

4. Farrell, R.H., and Stern, G.G., "A Measure of Tetrachoric Association," presented at APA convention, Sept. 1951, Chicago (unpublished; a copy of this paper may be obtained by writing The University of Chicago Computing Center).

5. Wright, B.D., "120X120 Principal Components Analysis: Symmetric (with Varimax Rotation and Plot)," Univ. Chicago Soc. Sci. Div. Computer Library (1963; unpublished).

DESCRIPTIVE TITLE	UNIVAC 1107 Linear Programming Package
CALLING NAME	LP1107
INSTALLATION NAME	The University of Notre Dame Computing Center
AUTHOR(S) AND AFFILIATION(S)	UNIVAC Division of Sperry Rand
LANGUAGE	Sleuth II/FORTRAN IV
COMPUTER	UNIVAC 1107
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Elizabeth Hutcheson, EIN Technical Representative, Computing Center, Univ. of Notre Dame, Notre Dame, Ind. 46556 Tel.: (219) 283-7784

FUNCTIONAL ABSTRACT

LP1107 is a generalized program for the solution of linear programming problems. More specifically, the LP1107 command structure is a mathematical programming control language; used to prepare a control sequence for a specific programming job.

Either the dual or simplex algorithm, or both, can be used. LP1107 incorporates a true programming language with logical capabilities and full macro capabilities. If a group of commands is being used repetitively, the user may incorporate it into his macro command library and issue the macro command instead. Single and double precision arithmetic are included.

REFERENCES

UNIVAC Systems Programming Library Services, *UP3897 1107LP User's Manual (preliminary)*, (UNIVAC, Sperry Rand Building, New York, 1964).

Aronofsky, J.S., "Growing Applications of Linear Programming with a View Towards Understanding Principles on Which Future LP Systems Must be Based," *Communications of the ACM*, (June 1964).

Copies of these references are available through the local UNIVAC representatives or from the EIN Office at the cost of reproduction and mailing.

DESCRIPTIVE TITLE	UNIVAC 1107 PERT/COST, PERT/TIME System
CALLING NAME	PERT (library tape)
INSTALLATION NAME	The University of Notre Dame Computing Center
AUTHOR(S) AND AFFILIATION(S)	UNIVAC Division of Sperry Rand Corporation
LANGUAGE	SLEUTH II and COBOL
COMPUTER	UNIVAC 1107
PROGRAM AVAILABILITY	Listings available
CONTACT	Elizabeth Hutcheson, EIN Technical Representative, Computing Center, University of Notre Dame, Notre Dame, Ind. 46556 Tel. (219) 283-7784

FUNCTIONAL ABSTRACT

The UNIVAC 1107 PERT System provides for integrated time and cost planning and control of research and development programs through implementation of the "work package" costing concept of the Department of Defense/National Aeronautics and Space Administration. The system is composed of two major program modules, PERT/COST and PERT/TIME, which generate reports for the project being analyzed. These reports include activity and event reports, a Work package/Activity listing, a Charge number analysis, etc., from PERT/TIME, and reports on the structure of labor and material costs from PERT/COST.

REFERENCES

PERT/COST General Reference Manual, (UP-4013), (UNIVAC Division of Sperry Rand Corp., N.Y.).

UNIVAC 1107/1108 PERT and CPM: Techniques in the Project Management, (UP 3952), (UNIVAC Division of Sperry Rand Corp., N.Y.).

Available through the local UNIVAC representative or through the EIN Office at the cost of reproduction and mailing.

000 0059

DESCRIPTIVE TITLE Basic Information Retrieval System

CALLING NAME BIRS

INSTALLATION NAME Michigan State University
Information Systems Laboratory

AUTHOR(S) AND AFFILIATION(S) John F. Vinsonhaler, Ph.D.
John M. Hafterson
Stuart W. Thomas, Jr.
Michigan State University

LANGUAGE USASI Full FORTRAN

COMPUTER CDC 3600, CDC 6500, CDC 6600
IBM 360 G-level
GE 600 Series, others

PROGRAM AVAILABILITY Decks and listings are currently available at cost from Michigan State University for non-profit institutions.

The program is distributed to profit-making institutions by [illegible] Technologies, 65 Whitney [illegible], Westport, Conn.

Maintenance for all users is provided by [illegible] Technologies.

CONTACT Dr. John F. Vinsonhaler [illegible]
Information Systems [illegible]
309 Computer Center, [illegible]
University, [illegible] 48823
Tel.: [illegible]

FUNCTIONAL ABSTRACT

BIRS is a general purpose system [illegible] sciences and education. [illegible] mental program modules designed [illegible] to use their own locally [illegible] tain a variety of information [illegible] and index creation are [illegible] may be viewed as a set of [illegible] may use these tools to [illegible] which best meets his [illegible]

It is not economically practical for individual educators and social scientists to develop their own special purpose programming system; they must share the costs of system developments by exchanging programs. General purpose systems like BIRS (with machine, data, and application independent programming) are ideally suited for free exchange among computer users.

REFERENCES

Vinsonhaler, J.F., Ed., *Technical Manual for the Basic Indexing and Retrieval System, BIRS 2.0,* Michigan State University, East Lansing, Mich., 1968.

Vinsonhaler, J.F., and Hafterson, J.M., Eds., *Technical Manual for the Basic Indexing and Retrieval System, Appendix I, BIRS* 2.5, Michigan State University, East Lansing, Mich., 1969.

Thomas, Jr., S.W., and Vinsonhaler, J.F., Eds., *Technical Manual for the Basic Information Analysis System, BIRS Technical Manual, Appendix II, BIAS 1.0,* Michigan State University, East Lansing, Mich., 1969.

Documentation is available at cost from the Information Systems Laboratory and from Hygain Technologies.

DESCRIPTIVE TITLE	FORTRAN Program to Assist in the Process of Political Reapportionment
CALLING NAME	T8 FSU BELOW
INSTALLATION NAME	Computer Center The Florida State University
AUTHOR(S) AND AFFILIATION(S)	William Below, Consultant Assembly Committee on Elections and Reapportionment California Legislature
LANGUAGE	CDC FORTRAN IV
COMPUTER	CDC 6400-65K
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Raymond Soller, Program Librarian/EIN Techncial Representative, Computing Center, Florida State University, Tallahassee, Fla. 32306 Tel.: (904) 599-4770

FUNCTIONAL ABSTRACT

BELOW is a program to assist with the development of political reapportionment plans using a set of procedures representative of the methods employed by legislatures to reapportion themselves. The actual algorithms utilized by the program are modifications of the computational techniques originally outlined by Stuart Nagel.[1]

The programmed procedures are those which follow. An initial district plan is determined by assigning geographic units to districts. This initial plan is fed into the processor along with demographic and political data for each unit. The program then tests each possible move of a unit from one district to another and then each trade of units between adjacent districts against a set of criteria. If the plan is approved according to the criterion; then the move or trade is made permanent. Where no more useful moves or trades can be made, the results are printed out and the program is terminated.

continued

The criterion used by the program is expressed as a linear combination of three quantities, SVEA, SVCA, SVPA;

$$\text{CRIT} = \text{WE} \cdot \text{SVEA} + \text{WC} \cdot \text{SVCA} + \text{SVPA}$$

where SVEA, SVCA, and SVPA are measures of population equality, compactness, and conformance to political goals, and WE and WC are weighting coefficients for the purpose of establishing the relative importance of the three quantities. CRIT, the criterion, is recomputed for each tentative move or trade. If its value is less than its previous value, the move or trade is consummated.

Population Equality

The measure of population equality is given by

$$\text{SVEA} = \sum_{J=1}^{\text{NDA}} \left[\frac{[\text{PD (J)} - \text{AVGPOP (J)}]}{\text{AVGPOP (J)}} \cdot 100 \right]^2$$

where NDA is the number of districts in the area, PD (J) is the population of district J and AVGPOP (J) is the ideal population. SVEA may be described as the sum of the squared fractional deviations from the ideal population for all districts in the area.

Multiplication of the fractional deviation by 100 is not significant to the operation of the program; it simply indicates that the deviation is output as a percentage. The use of the fractional rather than the absolute deviation is significant because it makes the effect of WE, (the relative importance of population equality), independent of the ideal population. Similar considerations apply to the decision to use total rather than average deviation for the area. If average deviation were used, then the proper setting of WE would depend on the number of districts in the area.

The opportunity for the user to set the ideal population for each district independently may at first seem arbitrary. In practice, this allows the program to be used on areas which contain portions of districts as well as whole districts.

This method of measuring population equality does not differ markedly from methods used by others in the field. One obvious alternative, namely, use of the sum of the absolute deviations, might be quite acceptable as a measure of equality to a student government, but it suffers from some computational disadvantages.

continued

As a population unit moves from one district to another, the absolute deviation changes only if one district is above, while the other is below the ideal population. The formula given here for SVEA however, has the convenient property that when two districts are both above or below the ideal population, a unit moved from one district to the other causes a change in SVEA, proportional to the difference in population between the two districts.

Compactness

The method originally tried for measuring compactness determines the "population moment" around the population center of the district. This method is attractive because it tends to minimize travel distance within the district and keep concentrations of population unbroken by district lines.

If there is a large number of units per district, the population moment method, unfortunately, can yield districts which are numerically compact, yet have very irregular boundaries. Since a legislature, as it reapportions itself, prefers to avoid even the appearance of gerrymandering, the author found the population moment method unsatisfactory for his purposes.

The alternative method, now in use, involves estimation of the perimeter of each district. The estimation is inherently inexact, but tends to produce district lines which are politically acceptable.

The program counts the perimeter elements in the perimeter of each district. A perimeter element is that portion of a unit perimeter which is shared with one other unit. Of course, these elements vary in length, but the program assumes each to have unit length.

One of the attractive features of the perimeter method is that it dispenses with the use of the X and Y coordinates required for the population moment method. What may be unattractive to some users is that the perimeter method does not tend to improve or even affect the overall shape of a district unless the initial district plan deviates widely from the desired characteristics.

Thus, if the program is given a district which is rectangular, and which is much longer than it is wide, the lines will be kept straight, but there will be little or no tendency for that district to change toward a square or circular shape. This property is often quite convenient in the legislative environment.

continued

In principle, it would be possible to measure the length of each perimeter element, include the data in the information fed to the processor, and work with a true perimeter measurement. Making those measurements, however, would be even more tedious than preparing the X and Y data, and may not necessarily improve program performance.

Political Considerations

The political portion of the criterion, CRIT, is given by

$$\mathrm{SVPA} = \sum_{J=1}^{NDA} \mathrm{WP}\ (J) \left[\frac{\mathrm{PAD}\ (J) \cdot 100}{\mathrm{PBD}\ (J)} - \mathrm{DESPR}\ (J) \right]^2$$

where WP (J) is a political weighting coefficient that may be set independently for each district, PAD (J) and PBD (J) are political quantities for district and DESPR (J) is the desired ratio of PAD (J) to PBD (J) expressed as a percentage.

PAD (J) and PDB (J) may stand for different things in different districts according to a number called MODE (J). If MODE (J) were set to 1, for example, and DESPR (J) to 55, then the program would set PAD (J) to equal registered Democrats and PBD (J) to equal total registered voters, and the goal for Democratic registration in that district would be 55%.

In the same manner, setting MODE (J) to 2 or 3 will establish a goal for the percentage of Negroes or persons with a Spanish surname in the population. The information for each unit necessary to establish any of the three proportions is carried in the memory of the processor. Consequently, there are no restrictions on moving a unit between districts with different modes.

Constraints on Moving and Trading

By the inclusion of the one card per unit in the data deck, the user may forbid the program to move or trade any number of units out of their original districts. The most common use of this feature is to ensure that incumbents are left in their own districts. Another use is to protect the integrity of municipal areas.

REFERENCES

1. Nagel, S., "Simplified Bipartisan Computer Redistricting," *Stanford Law Review*, 17, No. 5, May 1965.

2. Below, W., "The Computer as an Aid to Legislative Reapportionment," (Report to California Legislature, Assembly Committee on Elections and Reapportionment).

DESCRIPTIVE TITLE	TEXT360: A System for Producing Manuals
CALLING NAME	TEXT360
INSTALLATION NAME	Washington University Computing Facilities
AUTHOR(S) AND AFFILIATION(S)	S.L. Reed, International Business Machines
LANGUAGE	PL/I and Assembly Language
COMPUTER	IBM/360 under OS
PROGRAM AVAILABILITY	Proprietary; available for use but not for distribution
CONTACT	J. Philip Miller, Computing Facilities, Box 1152, Washington University, St. Louis, Mo. 63130 Tel.: (314) 863-0100 ext. 3141

FUNCTIONAL ABSTRACT

TEXT360 is a text-processing system with data-entry, data-updating, and page-formatting capabilities. The system, which runs under Operating System/360, consists of a main processor and several peripheral programs. Input to the system is free-form and is produced on the IBM 029 Card Punch. Output is camera-ready and is produced on the IBM 1403 Printer. The formatting capabilities of TEXT360 permit the insertion, deletion, and replacement of characters, words, lines, and groups of lines. In addition, blocks of text can be moved from one part of a document to another. One-column and two-column page format can be produced. Routine functions include hyphenation, line justification, column heading, and indentations. More complex functions include the generation of horizontal and vertical ruling for tables and figures. The program also allows the user to specify that related material (e.g., a table) is to be kept together, i.e., not split between columns or pages.

The four phases of the TEXT360 Formatting Processor, the TEXT360 Spelling Dictionary Update Program, the TEXT90-to-TEXT360 Master File Conversion Program and the Print/Punch Utility Program are

continued

written in the PL/1 language, supplemented by four small assembler-language routines used for character-set mapping. The TEXT360 Prescan and Peripheral Print Programs are written in System/360 assembler language.

TEXT360 is essentially the program used by IBM for the production of the SRL manuals.

REFERENCES

Reed, S.L., *TEXT360*, Contributed Program Library Documentation, Program 360D-29.4.001 (White Plains, N.Y.: IBM, 1967).

DESCRIPTIVE TITLE	Student Scheduling System
CALLING NAME	SCHEDULE
INSTALLATION NAME	The Michigan State University Computer Laboratory, Applications Programming
AUTHOR(S) AND AFFILIATION(S)	Withheld by request
LANGUAGE	FORTRAN IV and COMPASS
COMPUTER	CDC 3600
PROGRAM AVAILABILITY	Available for service at Michigan State University only
CONTACT	Mr. Anders Johanson, 324 Computer Center, Michigan State University, East Lansing, Mich. 48823 Tel.: (517) 355-4684

FUNCTIONAL ABSTRACT

School scheduling has two major parts, master schedule building and sectioning. Master schedule building is the process of deciding the courses to be offered and their time, place, instructor and seating capacity. Sectioning is the assigning of students into the courses of the master schedule.

The Student Scheduling System provides the school with a few valuable aids for the master schedule building process, performs the sectioning process and prints the individual students' schedules and the class lists.

DESCRIPTIVE TITLE	Mathematical Programming System
CALLING NAME	MPS/360
INSTALLATION NAME	Washington University Computing Facilities
AUTHOR(S) AND AFFILIATION(S)	IBM Application Program
LANGUAGE	360 Assembly Language
COMPUTER	IBM 360/50
PROGRAM AVAILABILITY	Proprietary: available for use at Washington University. Available for distribution from IBM.
CONTACT	Dr. G. B. Drebes, Mgr., Scientific Data Processing, Computing Facilities, Box 1098, Washington University, St. Louis, Mo. 63130 Tel.: (314) 863-0100 ext. 3141

FUNCTIONAL ABSTRACT

Linear programming (LP) is a mathematical technique designed to analyze the potentialities of alternate business activities and to choose those that permit the best use of resources in the pursuit of a desirable objective. It has many uses. For example, it can analyze capital, raw materials, manpower, plant and storage facilities and then translate its findings into minimum costs and maximum profits for its user. It can be used to allocate, assign, schedule, select or evaluate whatever possibilities limited resources possess for different jobs; it can blend, mix, distribute, control, order, budget, bid, cut, trim, price, purchase, and plan; it can deduce the most profitable method of transporting goods from plant to warehouse to outlet.

MPS/360 is composed of a set of procedures, a subset of which deals only with LP. The strategy for solving an LP problem is the ordered execution of a series of these procedures. A second set of procedures handles separable programming problems.

The user conveys the proposed strategy to MPS via the MPS control language. The procedure call statement of the control

continued

language calls the LP procedures and transfers arguments to them. MPS control statements are preprocessed by the control program COMPILER.

A catalogued procedure is available for simple linear programming models, reducing the user's work to merely data arrangement.

REFERENCES

Mathematical Programming System/360, Application Description (IBM Manual GH20-0136-3), (White Plains, N.Y.: IBM Corp. Tech. Pub. Depart., 1969).

Mathematical Programming System/360 (360A-CO-14X), Version 2, Control Language User's Manual (IBM Manual GH20-0290-3), (White Plains, N.Y.: IBM Corp. Tech. Pub. Depart., 1969).

Mathematical Programming System/360 (360A-CO-14X), Version 2, Linear and Separable Programming User's Manual (IBM Manual GH20-0476-0), (White Plains, N.Y.: IBM Corp. Tech. Pub. Depart., 1969).

Mathematical Programming System/360 (360A-CO-14X), Message Manual (IBM Manual GH20-0603-0), (White Plains, N.Y.: IBM Corp. Tech. Pub. Depart., 1969).

Copies of these manuals can be obtained through IBM branch offices.

DESCRIPTIVE TITLE	A Primal-Dual Transportation Algorithm
CALLING NAME	TRANSPRT
INSTALLATION NAME	Washington University Computing Facilities
AUTHOR(S) AND AFFILIATION(S)	D.E. Burlingame Washington University Computing Facilities
LANGUAGE	FORTRAN IV
COMPUTER	IBM 360/50
PROGRAM AVAILABILITY	Deck and listing presently available
CONTACT	Dr. C.B. Drebes, Mgr., Scientific Data Processing, Computing Facilities, Box 1098, Washington University, St. Louis, Mo. 63130 Tel.: (314) 863-0100 ext. 3141

FUNCTIONAL ABSTRACT

TRANSPRT solves a standard transportation problem using the primal-dual transportation algorithm. This program solves a model in which the objective is to "transport" a single commodity from various origins to different destinations at a minimum total shipping cost. The availability at each origin, the demand at each destination, and the cost to ship one unit of the project from any origin to any destination are required inputs te this model. This model can be applied to certain other types of industrial or business problems that have nothing to do with shipping. Personnel assignment, machine assignment, product and inventory scheduling are a few such applications.

The program utilizes integer arithmetic and requires integer input. If the sum of the origin availabilities is not equal to the sum of the destination requirements, an artificial origin or destination is set up, with zero costs, to handle the excess. The program will handle up to 50 origins and 150 destinations, using 94K bytes of core storage.

REFERENCES

Hadley, G., *Linear Programming* (Reading, Mass.: Addison-Wesley Pub. Co., Inc., 1962).

Behavioral Sciences

DESCRIPTIVE TITLE	Mann-Whitney U Test
CALLING NAME	MANNWH NUCC085
INSTALLATION NAME	Vogelback Computing Center, Northwestern University
AUTHOR(S) AND AFFILIATION(S)	Alan Lupa, Department of Chemical Engineering, Northwestern University Dennis R. Goldenson, Department of Political Science, Northwestern University
LANGUAGE	CDC FORTRAN IV
COMPUTER	CDC 6400
PROGRAM AVAILABILITY	Deck and listing presently available
CONTACT	Lorraine Borman, EIN Technical Representative, Vogelback Computing Center, Northwestern University, 2129 Sheridan Road, Evanston, Ill. 60201. Tel.: (312) 492-3682

FUNCTIONAL ABSTRACT

This program calculates the Mann-Whitney U statistic and reports its significance at one of five levels of confidence (that is, accept H_0, $P > .10$; reject H_0, $P < .10$, $P < .05$, $P < .02$, and $P < .01$). There is no limit to the number of problems that can be handled. Each subsample in a problem can have up to 2,000 cases. This is purely arbitrary; if one would want a larger sample-size capacity, he could merely alter the DIMENSION statements in the source deck to the limits of core storage. Printed output gives the problem number, alphanumeric problem label, the sample sizes, the input variable format, the median and decile range for each group, the U-test statistic (the smaller value), the value of Z where the larger group is greater than 8, and a statement of the level of confidence in the significance of U. (For a one-tailed test, the points of significance are found by dividing levels by two.) Since many behavioral-science data do not achieve interval scaling, the U test is a very useful alternative for one who does not wish to make the assumptions required by the parametric t test in determining if two independent samples are drawn from the same population. It is one of the most powerful of the nonparametric inferential statistics.

continued

References

1. Siegel, S., *Nonparametric Statistics for the Behavioral Sciences* (McGraw-Hill Book Co., Inc., New York, 1956).
2. Mann, H.B., and Whitney, D., "On a Test of Whether One of Two Random Variables is Stochastically Larger Than the Other," Ann. Math. Statistics *18*, 50-60 (1947).

DESCRIPTIVE TITLE	Subtle, Unbiased, Zealous Yatagen of Questionnaires
CALLING NAME	SUZYQ NUCC150
INSTALLATION NAME	Vogelback Computing Center, Northwestern University
AUTHOR(S) AND AFFILIATION(S)	Program written by Brent M. Rutherford, Northwestern University Converted for the CDC 6400 by Janos B. Koplyay
LANGUAGE	CDC FORTRAN IV
COMPUTER	CDC 6400 (Scope 3.1 O.S.)
PROGRAM AVAILABILITY	Deck and listing presently available
CONTACT	Lorraine Borman, EIN Technical Representative, Vogelback Computing Center, Northwestern University, 2129 Sheridan Road, Evanston, Ill. 60201 Tel.: (312) 492-3682

FUNCTIONAL ABSTRACT

This cycling program makes possible a scoring and weighting of questionnaire items, as well as providing comprehensive item-analysis measures, test-analysis measures, and factor analysis with varimax rotation of the tetrachoric interitem correlation matrix. For the purpose of item analysis, internal as well as external criterion scores may be utilized.

The program permits the user to correct and/or weight a set of test responses (including data that have already been scored). For the cases where some data values may be skipped or missing, there is an option for the insertion of a specified average value in their place. SUZYQ provides a count of the total-response score per item and per subject, expressed in *raw* score units, *z* score units, and *t* score units. In addition, a total-test-score frequency distribution is constructed, using the same three types of units.

The *total* test score is then subjected to an item analysis with corrections for the nonindependence of the calculated correlations.

continued

Finally, a two-way variance analysis is performed, using the method of Hoyt, with basic analytic parameters computed from the elements of the summary table. Options include item analysis by criterion and a derivation of up to ten factors from the tetrachoric interitem correlation matrix (using a Kaiser varimax factor analysis).

DESCRIPTIVE TITLE	FORTRAN Program for Guttman and Other Scalogram Analyses
CALLING NAME	GUTTSCL NUCC115
INSTALLATION NAME	Vogelback Computing Center, Northwestern University
AUTHOR(S) AND AFFILIATION(S)	Roland Werner, Department of Sociology, Syracuse University Donald G. Morrison, Northwestern University
LANGUAGE	CDC FORTRAN IV
COMPUTER	CDC 6400
PROGRAM AVAILABILITY	Deck and listing presently available
CONTACT	Lorraine Borman, EIN Technical Representative, Vogelback Computing Center, Northwestern University, 2129 Sheridan Road, Evanston, Ill. 60201 Tel.: (312) 492-3682

FUNCTIONAL ABSTRACT

GUTTSCL is designed to perform Guttman and other useful analyses on data supplied in scalogram form. Originally limited to 10 items with up to 7 response categories per item and 200 respondents, the capacity has recently been expanded to 30 items and 350 persons, the Chilton and Raju indices have been added, and other minor modifications have been made. Each data set constitutes an analysis. An unlimited number of such analyses can be treated serially.

Using the Cornell technique for scalogram analysis,[1] a scalogram is displayed in the output with the frequency of the response type, the Guttman and psychometric scores of the response type, and the number of errors in the response type. The item marginal frequencies and the item marginal probabilities are also provided. This information is utilized to calculate the Guttman coefficient of reproducibility. Other coefficients of reproducibility are based upon Loevinger's,[2] Sagi's,[3] and Green's[4] methods of counting errors in a response pattern. The expected coefficients of reproducibility

continued

for each of the error-counting methods are also calculated and are based on Goodman's[5] statistical techniques.

Since a wide range of flexibility in analyzing attitude data is desirable, additional indexes are calculated. These indexes are the Loevinger homogeneity index,[2] the Kuder–Richardson formula, the corrected Kuder–Richardson formula for test reliability,[6] the Green index of consistency,[4] the Borgatta error ratio,[7] the Menzel coefficient of scalability,[8] and the Schuessler χ^2 tests[9] for the frequency distribution of response types. The source of these indexes and their method of calculation is fully described in the references cited below.

REFERENCES

1. Guttman, L., "The Cornell Technique for Scale and Intensity Analysis," Educ. Psychol. Meas. *7*, 247-280 (1947).

2. Loevinger, J., "The Technique of Homogeneous Test Compared with Some Aspects of 'Scalogram Analysis' and Factor Analysis," Psychol. Bull. *45*, 507-529 (1948).

3. Sagi, P.C., "A Statistical Test for the Significance of a Coefficient of Reproducibility," Psychometrika *24*, 19-27 (1959).

4. Green, B.F., "A Method of Scalogram Analysis Using Summary Statistics," Psychometrika *21*, 79-88 (1956).

5. Goodman, L.A., "Simple Statistical Methods for Scalogram Analysis," Psychometrika *24*, 29-43 (1959).

6. Horst, P., "Correcting the Kuder–Richardson Reliability for Dispersion of Item Difficulty," Psychol. Bull. *50*, 371-374 (1953).

7. Borgatta, E.F., "An Error Ratio for Scalogram Analysis," Public Opinion Quart. *7*, 96-100 (1955).

8. Menzel, H., "A New Coefficient for Scalogram Analysis," Public Opinion Quart. *17*, 269-280 (1953).

9. Schuessler, K.F., "A Note on Statistical Significance of Scalograms," Sociometry *24*, 312-318 (1961).

continued

Guttman, L., "The Basis for Scalogram Analysis," in *Measurement and Prediction,* S.A. Stauffer, L. Guttman, and E.A. Suchman, *et al.*, Eds. (Princeton University Press, Princeton, N.J., 1950), pp. 60-90.

Werner, R., "A FORTRAN Program for Guttman and Other Scalogram Analyses," Syracuse Univ. (CPA 257). This manual provides the program listing and the instructions for the program's use. The control of the program is achieved with control cards that allow two modes of operation: one mode provides the user with summary frequencies of the attitude data; the other, in addition, calculates various indexes appropriate to attitude data.

DESCRIPTIVE TITLE	Automatic Interaction Detector
CALLING NAME	AID NUCC113
INSTALLATION NAME	Vogelback Computing Center, Northwestern University
AUTHOR(S) AND AFFILIATION(S)	J. Sonquist The University of Michigan
LANGUAGE	CDC FORTRAN IV
COMPUTER	CDC 6400
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Lorraine Borman, EIN Technical Representative, Vogelback Computing Center, Northwestern University, 2129 Sheridan Road Evanston, Ill. 60201 Tel.: (312) 492-3682

FUNCTIONAL ABSTRACT

AID was originally programmed for the IBM 7090 in the MAD language at The University of Michigan. A detailed discussion of the theory, methods, and control parameters of the program are contained in Ref. 1. Since 1964, two parameters have been added to the main-parameter card.

AID is focused on a particular kind of data-analysis problem, characteristic of many social-science research situations, in which the purpose of the analysis involves more than the reporting of descriptive statistics but may not necessarily involve the exact testing of specific hypotheses. In this type of situations the problem is often one of determining which of the variables, for which data have been collected, are related to the phenonmenon in question, under what conditions, and through what intervening processes, with appropriate controls for spuriousness.

AID is useful in studying the interrelationships among a set of up to 37 variables. Regarding one of the variables as a dependent variable, the analysis employs a nonsymmetrical branching process, based on variance-analysis techniques, to subdivide the sample into a series of subgroups that maximizes one's ability to predict values of the dependent variable. Linearity and additivity

continued

assumptions inherent in conventional multiple-regression techniques are not required. AID will handle variables that are only nominal scales, i.e., mere classifications.

REFERENCES

1. Morgan, J.N., and Sonquist, J.A., "The Detection of Interaction Effects: A Report on a Computer Program for the Selection of Optimal Combinations of Explanatory Variables," Univ. Mich. Inst. Soc. Res. Survey Res. Ctr. Monograph No. 35 (1964); copies may be obtained from the EIN office for the cost of duplication and mailing.

DESCRIPTIVE TITLE	Northwestern University Cross Classification and Tabulation
CALLING NAME	NUCROS NUCC017
INSTALLATION NAME	Vogelback Computing Center, Northwestern University
AUTHOR(S) AND AFFILIATION(S)	Betty Benson Vogelback Computing Center, Northwestern University
LANGUAGE	CDC FORTRAN IV
COMPUTER	CDC 6400
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Lorraine Borman, EIN Technical Representative, Vogelback Computing Center, Northwestern University, 2129 Sheridan Road, Evanston, Ill. 60201 Tel.: (312) 492-3682

FUNCTIONAL ABSTRACT

NUCROS is a general program for preparing crossclassifications (also called *crosstabulations* or *contingency tables*) in two, three, or four dimensions. For all tables, it is also possible to obtain row and column percentages, chi square, Kendall's tau, the Goodman–Kruskal gamma, and Somers D. A maximum of 99 problems may be executed in one run; each run is able to produce up to 72 tables from a given set of Control Cards and data. Data input may be from cards or tape. For each problem, a maximum of 9999 cases, with up to 80 variables/case, may be processed. The input data must be of the integer type; however, they may be recoded.

REFERENCES

Janda, K., *Data Processing* (Northwestern University Press, Evanston, Ill., 1968), 2nd ed.

Bonato, R. R., and Waxman, B. D., "A General Cross-Classification Program for Digital Computers," Behavioral Sci. *6*, No. 4, 347–357 (Oct. 1961).

DESCRIPTIVE TITLE	Test Scorer and Statistical Analysis
CALLING NAME	TSSA NUCC072
INSTALLATION NAME	Vogelback Computing Center Northwestern University
AUTHOR(S) AND AFFILIATION(S)	Richard Wolf and Leopold Klopfer The University of Chicago K. Jones Harvard University Graduate School of Education A. Gasche, B. Wright, and C. Bradford The University of Chicago
LANGUAGE	CDC FORTRAN IV
COMPUTER	CDC 6400
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Lorraine Borman, EIN Technical Representative, Vogelback Computing Center, Northwestern University, 2129 Sheridan Road, Evanston, Ill. 60201 Tel.: (312) 492-3682

FUNCTIONAL ABSTRACT

TSSA scores multiple-response tests that have a single correct response for each item, computes test and item statistics, computes the tetrachoric interitem correlation matrix, and performs a factor analysis and varimax rotation. The point-biserial correlation of each item with a criterion score also may be obtained. The number of alternative responses may vary from item to item.

Multiple-scoring keys may be used with the same set of data cards. This feature makes TSSA applicable for scoring and analysis of the Kuder preference record and similarly constructed multikeyed instruments. In addition, multiple jobs may be run.

Computational results include the following.

Individual Scores, including raw scores and scores corrected for guessing

continued

Item-Response Information, including the proportion of subjects selecting the correct response for each item (i.e., difficulty) and the proportion of subjects selecting each response

Test Statistics,[1] including the mean, standard deviation, skewness, and kurtosis (and their standard errors computed from formulas outlined by Fisher); test reliability (computed by Kuder—Richardson formula 20); a validity coefficient; and a Pearson product—moment correlation

Item-Analysis Information,[2,3] including the proportion of subjects passing an item, the item standard deviation, point-biserial correlations, reliability index, and validity index (computed from formulas given by Guilford and Gullicksen)

Tetrachoric Interitem Correlations[4]

Factor Analysis of Interitem Correlation Matrix,[5] with varimax rotation and plot

Multiple-Scoring Keys

A special feature of TSSA is the use of Multiple-Scoring Keys. Such a feature has two uses. (1) It allows a single set of responses to be scored in more than one way; practical application of this is the use of the program with data from an instrument such as the Kuder preference record where several scores, based on analyses of the same set of items, are desired. (2) A more common situation is the use of the program to obtain subtest scores and a total test score from one instrument. An application of this might be to score all the responses on a reading test and to obtain scores for reading speed, reading vocabulary, and level of comprehension, as well as a total reading score.

Multiple Jobs

The limitation of the procedures in the analyses of subtests is that only the raw scores for each individual or each subject will be meaningful. Corrected scores will *not* have any meaning. Thus, if an analysis excludes any items initially read, no corrected score will be printed. The reason for this is that the program treats *all* items excluded from analysis as incorrect responses. Thus, the use of a formula for correcting scores for guessing is fallacious in this instance. If the user desires meaningful corrected scores for subtests, he will have to reproduce his Data Cards and run multiple jobs, selecting for reading on any one job only those items for which analysis is desired, i.e., the subtest.

continued

The use of the Multiple-Jobs feature for scoring subtests yields meaningful corrected scores, as already mentioned. However, the advantage of the Multiple-Scoring-Key feature of the program is that the data need be read from cards only once. Subsequent reading of the data is done from binary tape (tape B3), thus achieving savings of time and, hence, cost.

The Multiple-Jobs feature may also be used for scoring and analyzing severly entirely different tests, with only a single loading of the program.

REFERENCES

1. Fisher, R.A., *Statistical Methods for Research Workers* (Oliver & Boyd, Edinburgh, 1954), 12th ed., pp. 70–75.

2. Guilford, J.P., *Psychometric Methods* (McGraw-Hill Book Co., Inc., New York, 1954), pp. 373–464.

3. Gullicksen, H., *Theory of Mental Tests* (John Wiley & Sons, Inc., New York, 1950), pp. 363–396.

4. Farrell, R.H., and Stern, G.G., "A Measure of Tetrachoric Association," presented at APA convention, Sept. 1951, Chicago (unpublished; a copy of this paper may be obtained by writing The University of Chicago Computing Center).

5. Wright, B.D., "120X120 Principle Components Analysis: Symmetric (with Varimax Rotation and Plot)," Univ. Chicago Soc. Sci. Div. Computer Library (1963; unpublished).

DESCRIPTIVE TITLE	Simulator of SAMOS (A Simple Imaginary Machine Language for Instruction)
CALLING NAME	SAMOS
INSTALLATION NAME	Computing Center, The Florida State University.
AUTHOR(S) AND AFFILIATION(S)	Marilyn Zupsich, Computing Center, The Florida State University
LANGUAGE	FORTRAN IV
COMPUTER	CDC 6400, 65K
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Marilyn Zupsich, Computing Center, The Florida State University, Tallahassee, Fla. 32306 Tel.: (904) 599-3418

FUNCTIONAL ABSTRACT

This program simulates the SAMOS computer described in the textbooks referred to below. The SAMOS programming language is a simplified machine-type language and is used to illustrate the logic of a computer to beginning-programming students. The simulator accepts programs written in the SAMOS language and executes them just as the mythical SAMOS computer would.

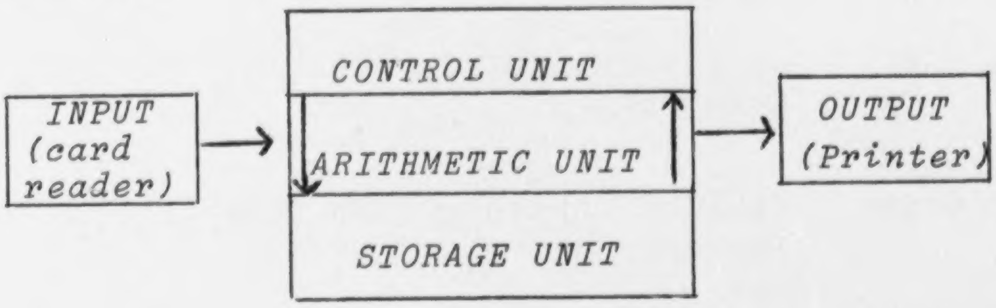
The SAMOS computer has 10,000 words of magnetic-core storage, each of which contains 10 characters and a sign (+ or -). Fifteen instruction types, covering arithmetic, branching, input, output, shifting, and the use of index registers, are simulated.

The basic design of the SAMOS computer is described in the following diagrams.

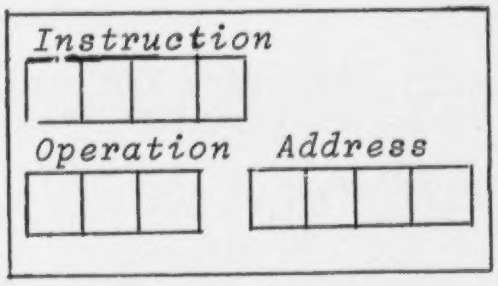
continued

Description of SAMOS Computer

Basic Design

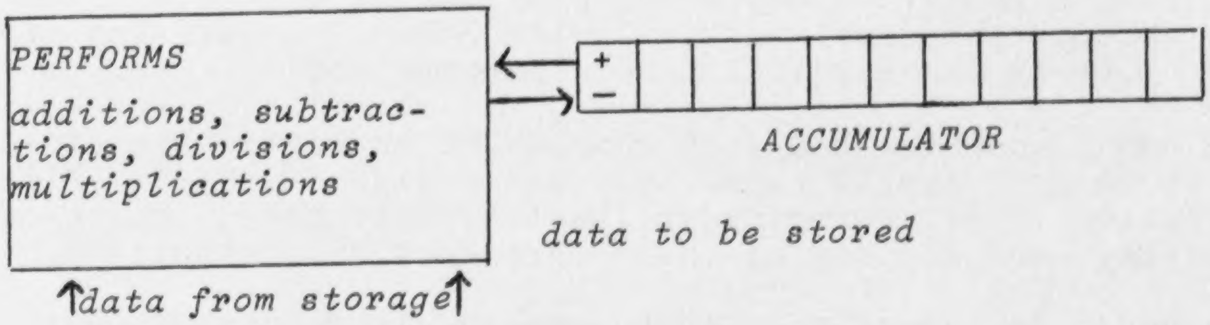

CONTROL UNIT
INPUT (card reader)
ARITHMETIC UNIT
OUTPUT (Printer)
STORAGE UNIT

Control Unit


Instruction
Operation
Address

The control unit is the part of the computer that determines which instruction is to be done next, decodes the operation to be done, and coordinates machine working with the programmed instruction. To begin, the instruction counter is set to show where the program begins.

Arithmetic Unit


PERFORMS
additions, subtractions, divisions, multiplications
+
−
ACCUMULATOR
data to be stored
data from storage

REFERENCES

1. School Mathematics Study Group, *Algorithms, Computation and Mathematics* [A.C. Vroman, Inc. (367 S. Pasadena Ave.), Pasadena, Calif. 91105, 1966].

2. Forsythe, A., Keenan, T., Organick, E., and Stenberg, W., *Computer Science: A First Course* (John Wiley & Sons, Inc., New York, 1969).

3. Forsythe, A., Keenan, T., Organick, E., and Stenberg, W., *Computer Science: A Primer* (John Wiley & Sons, Inc., New York, 1969).

DESCRIPTIVE TITLE	Synagraphic Computer-Mapping Program
CALLING NAME	SYMAP 5.12
INSTALLATION NAME	Computing Center, The Florida State University
AUTHOR(S) AND AFFILIATION(S)	Howard T. Fisher, Northwestern Technological Institute (overall design and mathematical model) Mrs. O.G. Benson, Northwestern University Computing Center (programming)
LANGUAGE	FORTRAN IV
COMPUTER	Originally developed on IBM 709; also on IBM 360/40 and IBM 7094 at Harvard University Available in EIN through The Florida State University CDC 6400
PROGRAM AVAILABILITY	Decks, listings, and documentation presently available from The Florida State University Computing Center
CONTACT	Ray Soller, Librarian, Computing Center, The Florida State University, Tallahassee, Fla. 32306 Tel.: (904) 599-3418

FUNCTIONAL ABSTRACT

The Synagraphic Computer-Mapping Program (SYMAP) produces maps that depict spatially disposed quantitative and qualitative information. Raw data of every kind (physical, social, economic, etc.) may be related, weighted, and aggregated in a graphic format by assigning values to the coordinate locations of data points or data zones. According to the application and desired representation of data, three basic types of mapping procedure may be specified: contour, conformant, or proximal.

CONTOUR—based on the use of *contour* lines, each of which represents a value remaining constant throughout its length. The map consists of closed curves that connect all points having the same numerical

continued

value. The value at each of the different levels of contour (where a single contour level will represent a particular data value) is determined by the program, according to the scale of the map and the range of the data. Between any two contour lines, a continuous variation is assumed. Therefore, the use of contour mapping should be restricted to the representation of continuously varying information, such as topography, rainfall, or population density.

CONFORMANT—based on the *conformance* to the boundaries of a data zone. This type of mapping is best suited to data for which the representation as a continuously varying surface is inappropriate owing to the significance of physical limits or boundaries. Each predefined data zone is assigned one data value and, depending on its numeric class (range), one representative character on the map itself. Local variation within the zone boundaries will not be apparent, but will, on the average, be correct.

PROXIMAL—based on *proximity* to a data point. In appearance, this type of map is similar to the conformant map. However, point information is used here to define the data zones. Each character location on the output map is assigned the value of the *nearest* data point, using nearest-neighbor techniques. Boundaries are then assumed along the lines where these values change. Then the mapping is carried out as in the conformant type.

REFERENCES

Robertson, J. C., "The SYMAP Programme for Computer Mapping," Cartographic 108–113 (Dec. 1967); taken from a report of the Select Committee on the Ordinance Survey of Scotland.

Fisher, M., "The Laboratory for Computer Graphics," Harvard Univ. Grad. School Design Suppl. (Summer 1967).

Shepard, D., "A Two-Dimensional Interpolation Function for Irregularly Spaced Data," Harvard Univ. Grad. School Design Lab. Computer Graphics (Feb. 1968); available from the Computing Center, The Florida State University. Deals with the subject of analyzing irregularly spaced data derived from a continuous surface. A method is developed for reconstructing the surface from the sampled data. This method is the main device used to generate the maps produced by SYMAP.

DESCRIPTIVE TITLE	Transportation Planning Package
CALLING NAME	TRAN/PLAN
INSTALLATION NAME	The Florida State University Computing Center
AUTHOR(S) AND AFFILIATION(S)	Charles E. Akers William T. Gaupin The Florida State University Computing Center
LANGUAGE	Primary Language: FORTRAN IV Secondary Language: COMPASS Assembly Language
COMPUTER	CDC 6400
PROGRAM AVAILABILITY	Programs generally available through the CDC users group. FSU will make them available to EIN members.
CONTACT	William T. Gaupin, Computing Center, The Florida State University, Tallahassee, Fla. 32306 Tel.: (904) 599-4770

FUNCTIONAL ABSTRACT

The Transportation Planning Package incorporates a set of analytical techniques that enable the user to estimate future transportation requirements and evaluate proposed systems. The collection of survey techniques, analysis method, and computer programs used by TRAN/PLAN have evolved over the past two decades through studies supported by State Highway Departments, U.S. Bureau of Public Roads, and the U.S. Department of Housing and Urban Development. The TRAN/PLAN programs perform the following functions, (1) process data accumulated from inventories or surveys of existing transportation facilities, (2) calibrate and analyze several parameters for a regional transportation network, and (3) predict future intraregional transportation demands. The accuracy of future estimates depends heavily upon the choice of a suitable model, effective intermediate analysis-of-program output, and iterative recalibration of model parameters. Reference 6 explains the overall purpose of transportation planning systems in greater detail.

continued

The TRAN/PLAN Package consists of the following subprograms[1].

(a)	RELINK	Relink nodes of network
(b)	NETWORK	Build and edit network
(c)	NETDVR	Network report generator
(d)	MINPATH	Build minimum paths
(e)	SKIM	Skim (summarize) tape of trees
(f)	COLLECT	Trip data collector
(g)	TRPBLDR	Build trip tables
(h)	ADDER	Add selected trip tables
(i)	SPLITER	Split trips by direction
(j)	UPDATE	Update trip volumes
(k)	MERGE	Merge trip tables
(l)	DEMERGE	Demerge trip tables
(m)	COMPRES	Compress or expand trip tables
(n)	SECTORS	Select sector of network
(o)	TRIPTAB	Trip-table summary generator
(p)	TRIPEND	Trip-end summary generator
(q)	GRAVITY	Gravity model
(r)	FRATAR	Fratar expansion
(s)	ASSIGN	Traffic assignment
(t)	CPACITY	Capacity restraint
(u)	MODSPLT	Modal split model
(v)	CAPRPT	Cross-reference historical-report generator
(w)	CPRSRPT	Historical-report generator
(x)	COMPARE	Statistical comparison of trip volumes
(y)	AVERAGE	Historical-record summary-report generator
(z)	TIMESUM	Time-table-report generator
(aa)	TRPFREQ	Trip frequency distribution

Restrictions which apply to all programs,

1. the maximum allowed zone (centroid) number is 650
2. the maximum allowed node number is 3000
3. the maximum number of tables (purposes) is 4
4. the maximum table number is 4
5. a stacked table may contain a maximum of 8 merged tables

REFERENCES

Akers, C.E. and Gaupin, W.T., "User Manual Transportation Programs for a CDC 6400 Computer," (The Florida State University Computing Center, November 1968; unpublished). Copies of this manual are available through the Librarian, The Florida State University Computing Center.

"Calibrating and Testing a Gravity Model for Any Size Urban Area," Bureau of Public Roads, Washington, D.C., October 1965.

continued

Manual Supplements for the Transportation Planning System, Highway and Local Government Department of South Austrailia, Adelaide, Austrailia, January 1966.

Martin, Brian V., "Minimum Path Algorithms for Transportation Planning," *Research Engineer*, December 1963.

"Modal Split," U.S. Department of Commerce, Bureau of Public Roads, Office of Planning, December 1966.

Shofer, R. and Goodyear, F., "Electronic Computer Applications in Urban Transportation Planning," Proceedings of the 22nd National Conference, Association for Computing Machinery Publication P–67, 1967.

Traffic Assignment Manual, Bureau of Public Roads, Washington, D.C., June 1964.

Transportation Planning System for the Control Data 3600 Computer, Data Centers Division of Control Data Corporation, Minneapolis, Minnesota, May 1965.

Fratar, T.J., "Vehicular Trip Distributions by Successvie Approximation," *Traffic Quarterly*, January 1954, pp. 53–65.

DESCRIPTIVE TITLE	A Package of Three FORTRAN Programs for Computerized Assistance in the Instruction of Beginning and Remedial Reading and the Evaluation of Such Instruction
CALLING NAME	DOVACK (a) DOVA; (b) DOVB; (c) REPORT
INSTALLATION NAME	Computer Center, The Florida State University
AUTHOR(S) AND AFFILIATION(S)	Florine Way, Project Director, ESEA Title III, Jefferson County, Florida, Board of Public Instruction
LANGUAGE	FORTRAN IV
COMPUTER	CDC 6400
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	William Petty, Computer Center, The Florida State University, Tallahassee, Fla. 32306 Tel.: (904) 599-3418

FUNCTIONAL ABSTRACT

DOVACK is a program package for accepting unstructured student dictations as input and returning these dictations in a structured form for study by the student of his own individual vocabulary. Each distinct English word is separated from the text of the dictation and listed alphabetically for comparison by the student of the word in and out of context. Each student's vocabulary is retained in a permanent file that is updated each time a new dictation is entered. Random-sample vocabulary-recognition tests may be prepared at the user's option from the individual student's vocabulary list. Punched-card information is provided at test time for use by the instructor in evaluating the success of the instruction. To facilitate the use of the program from the instructional site via remote terminals, the tasks are broken into three segments, each handled by a separate program.

(a) DOVA

...scans the unstructured dictations, echoes them with numbered

continued

lines, breaks the text into separate English words, alphabetizes and formats the word lists. The word lists, word frequency of occurrence, and location (line number) in the dictation are passed to DOVB via magnetic tape.

(b) DOVB

...merges each student's word list with the permanent files kept of the student's vocabulary. Two files are kept, one permanent (tape 1), the other semipermanent (tape 2). Random-sample vocabulary-recognition tests for the student group may be given from either file (or both). When a test is given from the semipermanent file, the file is erased, thus providing the capability of testing only over new words used since the last vocabulary test. The number of words on a test is determined by the formula N/[1.+(N–1.)/40.], where N is the total number of distinct words dictated. When tests are given, cards are prepared containing information on each student's vocabulary for use by REPORT.

(c) REPORT

...accepts the punched cards from DOVB, augmented by results of tests from the instructor and summarizes the instructional results in terms of new words learned by the student.

DESCRIPTIVE TITLE	Static Leontief Input—Output Analysis
CALLING NAME	INOUT
INSTALLATION NAME	The Pennsylvania State University Computation Center
AUTHOR(S) AND AFFILIATION(S)	M.C. Hallberg M. Swope Department of Agricultural Economics, The Pennsylvania State University
LANGUAGE	FORTRAN IV
COMPUTER	IBM System 360/67
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Daniel Bernitt, 105 Computer Building, The Pennsylvania State University, University Park, Pa. 16802 Tel.: (814) 865-9527

FUNCTIONAL ABSTRACT

INOUT will solve the static Leontief input—output problem for a model with as many as 175 sectors. From data consisting of a square input matrix, an output vector, and a set of final demand vectors, a great variety of results may be selectively opted, calculated, and printed. Among these are matrices of technical and interdependency coefficients, matrices of interdependency values and net effects, and some other useful vectors.

REFERENCES

Leontief, W., *et al.*, *Studies in the Structure of the American Economy*, Harvard Economic Research Proj. (Oxford University Press, New York, 1953); theoretical and empirical explorations in input—output analysis.

DESCRIPTIVE TITLE	Key Grapheme in Context
CALLING NAME	KGIC
INSTALLATION NAME	Washington University Computing Facilities
AUTHOR(S) AND AFFILIATION(S)	Joel Achtenberg Computing Facilities Washington University
LANGUAGE	PL/I
COMPUTER	S/360 under OS
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	J. Philip Miller, Computing Facilities, Washington University, St. Louis, Mo. 63130 Tel.: (314) 863-0100 ext. 4041

FUNCTIONAL ABSTRACT

The KGIC program was written to facilitate the analysis of the environmental distribution of graphic characters. It produces a KWIC-like listing of all occurrences of a given grapheme along with the graphic environment in which each instance appeared. The listing may be sorted either forward or backward from the key grapheme to facilitate inspection. Provision is made for specification by the user of special alphabets for foreign languages or for phonemic transcriptions.

If any word appears more than once in the data to be processed, only one set or records will be produced for that word. A counter will be increased and the frequency of occurrence printed in both the alphabetical listing and in the KGIC listing. The alpha listing thus will contain a complete frequency count of the corpus under consideration.

It should be noted that whenever the user can attribute phonetic or phonemic status to individual graphemes, the KGIC listing provides correspondingly significant information about phonetic and/or phonemic environments.

continued

The program produces the items listed below as output.

1. an alphabetical listing of all words processed by the KGIC program, along with their absolute and relative frequencies of occurrence,
2. the KGIC listing itself grouped by alphabetical character, with the absolute frequency of occurrence given for each unique occurrence, and the total number of occurrences and the total number of unique occurrences given for each character,
3. a summary table containing the absolute and relative frequencies of occurrence for both the total number of occurrences and for the total of unique occurences,
4. optionally, a horizontal bar graph of the relative frequencies of all occurrences of each grapheme,
5. a number of summary statistics, i.e.,
    - a) total number of words processed, i.e., tokens
    - b) average length of word
    - c) total number of unique words, i.e., types
    - d) the type/token ratio
    - e) total number of characters processed
    - f) total number of unique occurrences of all characters
6. a statement of all program options used in a particular run, and a complete listing of the EMICTT, defining the alphabet in use for that run.

000 0055

DESCRIPTIVE TITLE	Bio-Medical Multivariate Statistical Programs
CALLING NAME	BMD
INSTALLATION NAME	University of Notre Dame Computing Center
AUTHOR(S) AND AFFILIATION(S)	School of Medicine University of California, Los Angeles UNIVAC Division of Sperry Rand Corporation College of Business Administration and Computing Center University of Notre Dame
LANGUAGE	FORTRAN IV
COMPUTER	UNIVAC 1107
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Elizabeth Hutcheson, EIN Technical Representative, Computing Center, University of Notre Dame, Notre Dame, Ind. 46556 Tel.: (219) 283-7784

FUNCTIONAL ABSTRACT

The BMD system is a package of computer programs designed to do both basic data processing and the subsequent statistical analysis. The programs have been prepared in an easy-to-use parametric form so that the researcher may adapt them to a wide variety of statistical problems. For further details of the package, see the User Instructions. The BMD is available at the University of Notre Dame in its 1967 edition.

REFERENCES

Dixon, W.J. (Ed.), *BMD: Biomedical Computer Programs,* (Univ. of Calif. Press, Berkeley, 1967, 1970).

Univ. of Notre Dame Computing Center, *BMD for the UNIVAC 1107*, (Rough draft, 1967 ed. of BMD). Available from Univ. of Notre Dame Computing Center, Notre Dame, Ind.

000 0055

DESCRIPTIVE TITLE	Inquirer II System for Content Analysis
CALLING NAME	I/II
INSTALLATION NAME	Washington University Computing Facilities
AUTHOR(S) AND AFFILIATION(S)	J. Philip Miller Washington University Computing Facilities
LANGUAGE	PL/I
COMPUTER	S/360 under OS
PROGRAM AVAILABILITY	Source listings and complete documentation available from Documentation Librarian, Computing Facilities, Box 1152, Washington University, St. Louis, Mo. 63130
CONTACT	J. Philip Miller, Computing Facilities, Box 1152, Washington University, St. Louis, Mo. 63130 Tel.: (314) 863–0100 ext. 4041

FUNCTIONAL ABSTRACT

The Inquirer II is a set of computer programs comparable to, but more flexible than, its predecessor, the General Inquirer, developed by Stone and his colleagues at Harvard. The original version of the General Inquirer System was designed for problems encountered in the content analysis of textual and verbal data. The General Inquirer was implemented for the IBM 7090–7094 computer along with the IBM 1401 computer. A later version of the General Inquirer was designed by Psathas and Miller to be used only with an IBM 1401 computer with an IBM 1311 disk drive. Stone originally described the General Inquirer as "a set of computer programs to (a) identify, systematically within text, instances of words and phrases that belong to categories specified by the investigator; (b) count occurrences and specify co-occurrences of these categories; (c) print and graph tabulations; (d) perform statistical tests; (e) sort and regroup sentences according to whether they contain instances of a particular category of combination of categories." The Inquirer II contains these capabilities and also

continued

allows for more elaborate analysis of the data. The I/II is able to make more elaborate contextual searches of the data and provide more options to the potential users. These options are described in detail in the *Inquirer II Programmer's Guide*.

Content analysis may be defined as a research technique which includes a systematic identification of theoretically relevant constructs in textual data. Content analysis is usually performed so that inferences can be made about the source or originator of the message, the message itself or the intended receiver of the message. The investigator communicates the constructs and the rules by which they may be identified within the corpus of text by means of dictionary. This dictionary is either one of his own construction or one which has been utilized in previous research.

REFERENCES

Miller, J. Philip, editor, *Inquirer II Programmer's Guide*, 2nd Edition, 1970, Washington University, (Available from Documentation Librarian, Computing Facilities, Box 1152, Washington University, St. Louis, Mo. for $5.00)

Stone, P.J., Dunphy, D.C., Smith, M.S., and Ogilvie, D.M. The General Inquirer: *A Computer Approach to Content Analysis*, Cambridge, Mass., The MIT Press, 1966.

DESCRIPTIVE TITLE	Basic Information Retrieval System
CALLING NAME	BIRS
INSTALLATION NAME	Michigan State University Information Systems Laboratory
AUTHOR(S) AND AFFILIATION(S)	John F. Vinsonhaler, Ph.D. John M. Hafterson Stuart W. Thomas, Jr. Michigan State University
LANGUAGE	USASI Full FORTRAN
COMPUTER	CDC 3600, CDC 6500, CDC 6600 IBM 360 G-level GE 600 Series, others
PROGRAM AVAILABILITY	Decks and listings are currently available at cost from Michigan State University for non-profit institutions. The program is distributed to profit-making institutions by Hygain Technologies, 65 Whitney Street, Westport, Conn. Maintenance for all users is provided by Hygain Technologies.
CONTACT	Dr. John F. Vinsonhaler, Director, Information Systems Laboratory, 309 Computer Center, Michigan State University, East Lansing, Mich. 48823 Tel.: (517) 353-7284

FUNCTIONAL ABSTRACT

BIRS is a general purpose system of programs for the behavioral sciences and education. Essentially, BIRS is a set of fundamental program modules designed to allow scholars and scientists to use their own locally based computer to construct and maintain a variety of information systems. Search, maintenance, and index creation are performed automatically. Thus, BIRS may be viewed as a set of essential tools; the research worker may use these tools to construct the type of information system which best meets his immediate needs.

continued

It is not economically practical for individual educators and social scientists to develop their own special purpose programming system; they must share the costs of system developments by exchanging programs. General purpose systems like BIRS (with machine, data, and application independent programming) are ideally suited for free exchange among computer users.

REFERENCES

Vinsonhaler, J.F., Ed., *Technical Manual for the Basic Indexing and Retrieval System, BIRS 2.0*, Michigan State University, East Lansing, Mich., 1968.

Vinsonhaler, J.F., and Hafterson, J.M., Eds., *Technical Manual for the Basic Indexing and Retrieval System, Appendix I, BIRS 2.5*, Michigan State University, East Lansing, Mich., 1969.

Thomas, Jr., S.W., and Vinsonhaler, J.F., Eds., *Technical Manual for the Basic Information Analysis System, BIRS Technical Manual, Appendix II, BIAS 1.0*, Michigan State University, East Lansing, Mich., 1969.

Documentation is available at cost from the Information Systems Laboratory and from Hygain Technologies.

DESCRIPTIVE TITLE	Analysis of Change-over Experiments
CALLING NAME	ZFE-03, ZFE-04
INSTALLATION NAME	Office of Data Analysis Research Educational Testing Service (ETS)
AUTHOR(S) AND AFFILIATION(S)	Procedure: G. Beall formerly of Gillette Razor Co. Program: V. Halfmann formerly of ETS
LANGUAGE	FORTRAN
COMPUTER	IBM 360/65
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Mr. Ernest Anastasio, Off. of Data Anal. Research, Educational Testing Service, Rosedale Road, Princeton, N.J. 08540 Tel.: (609) 921-9000 ext. 2552

FUNCTIONAL ABSTRACT

The class of designs covered by this program is latin squares or Youden rectangles (incomplete latin squares). These may be repeated fully or in part. The design may be defective, i.e., certain whole rows may be missing, but no allowance has been made for missing cells, i.e., single observations.

The purpose of this program is primarily to analyze the results for effect of treatment with allowance for carry-over of preceding treatment. There is also direct testing of the significance of carry-over. There is included parallel estimation and testing of significance without allowance for carry-over. The program is self-contained and does not require any external subroutines, such as might be presumed to exist in one form or another at computation centers.

The program, as presently stored, allows analysis for designs up to 40 rows (or blocks). This is limited by the dimensions of the data matrices. It could be readily enough changed to 500 or 1000 if such an experiment were involved. The analysis has been contrived so that such change does not increase the size of the matrix involved in equation solving.

continued

Explanation

The basic equations are,

1. $y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_k + \varepsilon_{ijk}$

where y_{ijk} is an observation assumed built of a general level μ, effect of the i^{th} row or individual α_i, the j^{th} period or column β_j, the k^{th} treatment γ_k and extraneous variability ε_{ijk}. This equation obtains for the first column or period when there has been no conditioning period. For the following periods,

2. $y_{ijk\ell} = \mu + \alpha_i + \beta_j + \gamma_k + \delta_\ell + \varepsilon_{ijk\ell}$

where the effect of the ℓ^{th} treatment in the preceding period is δ_ℓ. For a conditioned experiment Equ. (2) obtains in all columns.

The data actually considered are the differences within rows such as,

$$y_{ijk\ell} - y_{i'j'k'\ell'} = \beta_j - \beta_{j'} + \gamma_k - \gamma_{k'} + \delta_\ell - \delta_{\ell'}$$

$$(j' \neq j \;,\; k' \neq k \;,\; \ell' \neq \ell) \;.$$

These differences are then set forth in a matrix. Thus for an unconditioned latin square for which the first line is,

Design:	(1)	(2)	(4)	(3)
Result:	4	5	7	6

we may consider the two differences,

$$y_{111} - y_{1221} = \beta_1 - \beta_2 + \gamma_1 - \gamma_2 - \delta_1 + \varepsilon_{111} - \varepsilon_{1221}$$

$$y_{1221} - y_{1342} = \beta_2 - \beta_3 + \gamma_2 - \gamma_4 + \delta_1 - \delta_2 + \varepsilon_{1221} - \varepsilon_{1342}$$

which results in two lines of the matrix as follows:

β_1	β_2	β_3	β_4	γ_1	γ_2	γ_3	γ_4	δ_1	δ_2	δ_3	δ_4	Result
+1	−1			+1	−1			−1				−1
	+1	−1			+1		−1	+1	−1			−2

Least-squares equations are in the same form. For instance, to get the equation associated with β_2, each line is multiplied by its content in the β_2 column and the product accumulated over all columns. For each set of effects (β, γ or δ), the last equation is replaced by a condition equation,

continued

$$\sum_j \hat{\beta}_j = \sum \hat{\gamma}_k = \sum_\ell \hat{\delta}_\ell = 0 .$$

The analysis without carry-over (SANS DELTA) is gotten by replacing, temporarily, all equations appropriate to δ by $\hat{\delta}_\ell = 0$.

The analysis without treatment (SANS GAMMA) is obtained by the temporary replacement, $\hat{\gamma}_k = 0$.

The residual variability is gotten in several steps. First the variability residual on β , γ and δ is

$$3. \quad \sum_{ijk} y^2_{ijk} + \sum_{ijk\ell} y^2_{ijk\ell} - \sum_j \hat{\beta}_j(\sum_{i} y_{ijk} + \sum_{ik\ell} y_{ijk\ell})$$

$$- \sum_k \hat{\gamma}_k(\sum_{ij} y_{ijk} + \sum_{ij\ell} y_{ijk\ell}) - \sum_\ell \hat{\delta}_\ell \sum_{ijk} y_{ijk\ell} .$$

Secondly, an estimate of $\hat{\mu}$ is made by finding from Equ. (1) or (2) the mean of the values

$$y'_{ijk} = y_{ijk} - \hat{\beta}_j - \hat{\gamma}_k$$

$$y'_{ijk\ell} = y_{ijk\ell} - \hat{\beta}_j - \hat{\gamma}_k - \hat{\delta}_\ell \quad .$$

Thirdly, estimates of $\hat{\alpha}_i$ are made by finding the mean of the values

$$y''_{ijk} = y_{ijk} - \hat{\mu}$$

$$y''_{ijk\ell} = y'_{ijk\ell} - \hat{\mu} .$$

The residual variability as from equation (3) is then further and finally reduced by

$$\hat{\mu} \sum_{ijk\ell} y_{ijk\ell} + \sum_i \alpha_i \sum_{jk\ell} y_{ijk\ell} .$$

Residual variability on the effects of $\hat{\mu}$, rows and columns only (SANS DELTA & SANS GAMMA) is gotten by the formula familiar in analysis of variance.

continued

The test of significance for treatments without allowance for carry-over is based on residual variability SANS DELTA less residual variability SANS DELTA & SANS GAMMA. The test with allowance for carry-over is from residual variability on FULL MATRIX less that on SANS GAMMA. The test for carry-over is from residual variability on FULL MATRIX less that on SANS DELTA.

It need only be added that there is incorporated a test on whether the situation is underdetermined. The program counts the number of different row patterns, multiples this number by the number of columns, and checks whether the result exceeds the number of independent parameters to be estimated. In the case of underdetermination, it refuses to analyze. A second type of refusal arises if the simultaneous equations prove insoluble, which may arise if the design is redundant. Finally, if there is no residual freedom, the program will estimate the parameters but declare F = 0.

DESCRIPTIVE TITLE	Multidimensional Scalogram Analysis
CALLING NAME	MSA-I
INSTALLATION NAME	Office of Data Analysis Research Educational Testing Service
AUTHOR(S) AND AFFILIATION(S)	Procedure due to: L. Guttman J. Lingoes University of Michigan Program due to: J. Lingoes University of Michigan Adaptation due to: D. Kirk Educational Testing Service Advisor on use: H. Harman Educational Testing Service
LANGUAGE	MAP
COMPUTER	IBM 360/65
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Mr. Ernest Anastasio, Off. of Data Anal. Research, Educational Testing Service, Rosedale Road, Princeton, N.J. 08540 Tel.: (609) 921-9000 ext. 2552

FUNCTIONAL ABSTRACT

MSA-I is a program to map types (individuals having the same profile over a set of variables or items) onto an Euclidean space with minimum dimensionality. No assumptions are required about the underlying distributions, the scaling properties of the items, or their ordering. The only requirement is that the categories of each item be mutually exclusive and exhaustive. Types are represented as points in space, each item is a partition of the space, and each category is a region. All types who fall in a given category of a particular item are constrained to be closer to boundary markers of the same category than to delimiters of other categories of the same item. The program we have is a modification of the MSA-I package as distributed

continued

by Dr. Lingoes. An option or modification provides additional plots for each category with the number inserted in the plots corresponding to the response to the item rather than the ID number as in the distributed version.

References and some copies of the substantial literature related to this problem are available through the contact person.

REFERENCE

Lingoes, J.C., "Multiple Scalogram Analysis—A Set-theoretic Model for Analyzing Dichotomous Items," J. of Ed. & Psych. Meas., *23*, pp. 501-524, (1963).

DESCRIPTIVE TITLE	FORTRAN Program to Assist in the Process of Political Reapportionment
CALLING NAME	T8 FSU BELOW
INSTALLATION NAME	Computer Center The Florida State University
AUTHOR(S) AND AFFILIATION(S)	William Below, Consultant Assembly Committee on Elections and Reapportionment California Legislature
LANGUAGE	CDC FORTRAN IV
COMPUTER	CDC 6400-65K
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Raymond Soller, Program Librarian/EIN Techncial Representative, Computing Center, Florida State University, Tallahassee, Fla. 32306 Tel.: (904) 599-4770

FUNCTIONAL ABSTRACT

BELOW is a program to assist with the development of political reapportionment plans using a set of procedures representative of the methods employed by legislatures to reapportion themselves. The actual algorithms utilized by the program are modifications of the computational techniques originally outlined by Stuart Nagel.[1]

The programmed procedures are those which follow. An initial district plan is determined by assigning geographic units to districts. This initial plan is fed into the processor along with demographic and political data for each unit. The program then tests each possible move of a unit from one district to another and then each trade of units between adjacent districts against a set of criteria. If the plan is approved according to the criterion; then the move or trade is made permanent. Where no more useful moves or trades can be made, the results are printed out and the program is terminated.

continued

The criterion used by the program is expressed as a linear combination of three quantities, SVEA, SVCA, SVPA;

$$\text{CRIT} = \text{WE} \cdot \text{SVEA} + \text{WC} \cdot \text{SVCA} + \text{SVPA}$$

where SVEA, SVCA, and SVPA are measures of population equality, compactness, and conformance to political goals, and WE and WC are weighting coefficients for the purpose of establishing the relative importance of the three quantities. CRIT, the criterion, is recomputed for each tentative move or trade. If its value is less than its previous value, the move or trade is consummated.

Population Equality

The measure of population equality is given by

$$\text{SVEA} = \sum_{J=1}^{\text{NDA}} \left[\frac{[\text{PD (J)} - \text{AVGPOP (J)}]}{\text{AVGPOP (J)}} \cdot 100 \right]^2$$

where NDA is the number of districts in the area, PD (J) is the population of district J and AVGPOP (J) is the ideal population. SVEA may be described as the sum of the squared fractional deviations from the ideal population for all districts in the area.

Multiplication of the fractional deviation by 100 is not significant to the operation of the program; it simply indicates that the deviation is output as a percentage. The use of the fractional rather than the absolute deviation is significant because it makes the effect of WE, (the relative importance of population equality), independent of the ideal population. Similar considerations apply to the decision to use total rather than average deviation for the area. If average deviation were used, then the proper setting of WE would depend on the number of districts in the area.

The opportunity for the user to set the ideal population for each district independently may at first seem arbitrary. In practice, this allows the program to be used on areas which contain portions of districts as well as whole districts.

This method of measuring population equality does not differ markedly from methods used by others in the field. One obvious alternative, namely, use of the sum of the absolute deviations, might be quite acceptable as a measure of equality to a student government, but it suffers from some computational disadvantages.

continued

As a population unit moves from one district to another, the absolute deviation changes only if one district is above, while the other is below the ideal population. The formula given here for SVEA however, has the convenient property that when two districts are both above or below the ideal population, a unit moved from one district to the other causes a change in SVEA, proportional to the difference in population between the two districts.

Compactness

The method originally tried for measuring compactness determines the "population moment" around the population center of the district. This method is attractive because it tends to minimize travel distance within the district and keep concentrations of population unbroken by district lines.

If there is a large number of units per district, the population moment method, unfortunately, can yield districts which are numerically compact, yet have very irregular boundaries. Since a legislature, as it reapportions itself, prefers to avoid even the appearance of gerrymandering, the author found the population moment method unsatisfactory for his purposes.

The alternative method, now in use, involves estimation of the perimeter of each district. The estimation is inherently inexact, but tends to produce district lines which are politically acceptable.

The program counts the perimeter elements in the perimeter of each district. A perimeter element is that portion of a unit perimeter which is shared with one other unit. Of course, these elements vary in length, but the program assumes each to have unit length.

One of the attractive features of the perimeter method is that it dispenses with the use of the X and Y coordinates required for the population moment method. What may be unattractive to some users is that the perimeter method does not tend to improve or even affect the overall shape of a district unless the initial district plan deviates widely from the desired characteristics.

Thus, if the program is given a district which is rectangular, and which is much longer than it is wide, the lines will be kept straight, but there will be little or no tendency for that district to change toward a square or circular shape. This property is often quite convenient in the legislative environment.

continued

In principle, it would be possible to measure the length of each perimeter element, include the data in the information fed to the processor, and work with a true perimeter measurement. Making those measurements, however, would be even more tedious than preparing the X and Y data, and may not necessarily improve program performance.

Political Considerations

The political portion of the criterion, CRIT, is given by

$$\text{SVPA} = \sum_{J=1}^{\text{NDA}} \text{WP (J)} \left[\frac{\text{PAD (J)} \cdot 100}{\text{PBD (J)}} - \text{DESPR (J)} \right]^2$$

where WP (J) is a political weighting coefficient that may be set independently for each district, PAD (J) and PBD (J) are political quantities for district and DESPR (J) is the desired ratio of PAD (J) to PBD (J) expressed as a percentage.

PAD (J) and PDB (J) may stand for different things in different districts according to a number called MODE (J). If MODE (J) were set to 1, for example, and DESPR (J) to 55, then the program would set PAD (J) to equal registered Democrats and PBD (J) to equal total registered voters, and the goal for Democratic registration in that district would be 55%.

In the same manner, setting MODE (J) to 2 or 3 will establish a goal for the percentage of Negroes or persons with a Spanish surname in the population. The information for each unit necessary to establish any of the three proportions is carried in the memory of the processor. Consequently, there are no restrictions on moving a unit between districts with different modes.

Constraints on Moving and Trading

By the inclusion of the one card per unit in the data deck, the user may forbid the program to move or trade any number of units out of their original districts. The most common use of this feature is to ensure that incumbents are left in their own districts. Another use is to protect the integrity of municipal areas.

REFERENCES

1. Nagel, S., "Simplified Bipartisan Computer Redistricting," *Stanford Law Review, 17*, No. 5, May 1965.

2. Below, W., "The Computer as an Aid to Legislative Reapportionment," (Report to California Legislature, Assembly Committee on Elections and Reapportionment).

DESCRIPTIVE TITLE	Multiple Scalogram Analysis	
CALLING NAME	MSA	
INSTALLATION NAME	Office of Computational Sciences Educational Testing Service	
AUTHOR(S) AND AFFILIATION(S)	Procedure due to:	L. Guttman J. Lingoes University of Michigan
	Program due to:	J. Lingoes University of Michigan
	Adaptation due to:	D. Kirk Educational Testing Service
	Advisor on use:	H. Harman Educational Testing Service
LANGUAGE	MAP	
COMPUTER	IBM 360/65	
PROGRAM AVAILABILITY	Decks and listings presently available	
CONTACT	Mr. Ernest Anastasio, Office of Data Analysis Research, Educational Testing Service, Rosedale Road, Princeton, N.J. 08540 Tel.: (609) 921-9000 ext. 2552	

FUNCTIONAL ABSTRACT

This program performs a multiple scalogram analysis using the method of James Lingoes[1]. Do not confuse this with Guttman-Lingoes Multidimensional scalogram analysis which is available under the name of MSA-I (EIN Abstract 000 0070).

REFERENCES

1. Lingoes, J.C., "A Set-Theoretic Model for Analyzing Dichotomous Items," J. of Ed. Psych. Meas., *23*, pp. 501–524, (1963).

DESCRIPTIVE TITLE	Synagraphic Mapping Program
CALLING NAME	SYMAP
INSTALLATION NAME	Washington University Computing Facilities
AUTHOR(S) AND AFFILIATION(S)	Harvard University Graduate School of Design
LANGUAGE	FORTRAN IV
COMPUTER	IBM 360/50
PROGRAM AVAILABILITY	Proprietary; available for use but not distribution
CONTACT	Dr. Charles Drebes, Manager, Scientific Data Processing, Computing Facilities, Box 1098, Washington University, St. Louis, Mo. 63130 Tel.: (314) 863-0100 ext. 3141

FUNCTIONAL ABSTRACT

SYMAP is a computer program designed to allow city planners, geographers and others to produce low cost graphic displays of spatial patterns using standard computer line printers, by producing maps which graphically depict spatially disposed quantitative and qualitative information. It is suited to a broad range of applications, and is provided with numerous options to meet widely varying requirements. Raw data of every kind (physical, social, economic, etc.) when given to the computer may be related, manipulated, weighted, and aggregated in any manner desired. By assigning values to the coordinate locations of data points or data zones, one or more of three types of map may be produced, as specified by the user: conformant (choropleth), contour, and proximal. Potential applications are independent of the scale at which one wishes to display data. Studies (at other universities) have included a living cell, land parcels, blocks, tracts, towns, states, and continents. In each case, a common factor was the spatial distribution of a variable and a need to display the patterns associated with this distribution.

REFERENCES

Reference Manual for Synagraphic Computer Mapping—"SYMAP" (Cambridge, Mass.: Harvard Univ. Grad. Sch. of Design, Comp. Graph. Lab.).

DESCRIPTIVE TITLE	Mathematical Programming System
CALLING NAME	MPS/360
INSTALLATION NAME	Washington University Computing Facilities
AUTHOR(S) AND AFFILIATION(S)	IBM Application Program
LANGUAGE	360 Assembly Language
COMPUTER	IBM 360/50
PROGRAM AVAILABILITY	Proprietary: available for use at Washington University. Available for distribution from IBM.
CONTACT	Dr. C. B. Drebes, Mgr., Scientific Data Processing, Computing Facilities, Box 1098, Washington University, St. Louis, Mo. 63130 Tel.: (314) 863-0100 ext. 3141

FUNCTIONAL ABSTRACT

Linear programming (LP) is a mathematical technique designed to analyze the potentialities of alternate business activities and to choose those that permit the best use of resources in the pursuit of a desirable objective. It has many uses. For example, it can analyze capital, raw materials, manpower, plant and storage facilities and then translate its findings into minimum costs and maximum profits for its user; it can be used to allocate, assign, schedule, select or evaluate whatever possibilities limited resources possess for different jobs; it can blend, mix, distribute, control, order, budget, bid, cut, trim, price, purchase, and plan; it can deduce the most profitable method of transporting goods from plant to warehouse to outlet.

MPS/360 is composed of a set of procedures, a subset of which deals only with LP. The strategy for solving an LP problem is the ordered execution of a series of these procedures. A second set of procedures handles separable programming problems.

The user conveys the proposed strategy to MPS via the MPS control language. The procedure call statement of the control

continued

language calls the LP procedures and transfers arguments to them. MPS control statements are preprocessed by the control program COMPILER.

A catalogued procedure is available for simple linear programming models, reducing the user's work to merely data arrangement.

REFERENCES

Mathematical Programming System/360, Application Description (IBM Manual GH20-0136-3), (White Plains, N.Y.: IBM Corp. Tech. Pub. Depart., 1969).

Mathematical Programming System/360 (360A-CO-14X), Version 2, Control Language User's Manual (IBM Manual GH20-0290-3), (White Plains, N.Y.: IBM Corp. Tech. Pub. Depart., 1969).

Mathematical Programming System/360 (360A-CO-14X), Version 2, Linear and Separable Programming User's Manual (IBM Manual GH20-0476-0), (White Plains, N.Y.: IBM Corp. Tech. Pub. Depart., 1969).

Mathematical Programming System/360 (360A-CO-14X), Message Manual (IBM Manual GH20-0603-0), (White Plains, N.Y.: IBM Corp. Tech. Pub. Depart., 1969).

Copies of these manuals can be obtained through IBM branch offices.

DESCRIPTIVE TITLE	University of Maryland Test Scoring Program (Version 4)
CALLING NAME	UOM 32
INSTALLATION NAME	University of Maryland Computer Science Center
AUTHOR(S) AND AFFILIATION(S)	Dr. George Green University of Maryland
LANGUAGE	FORTRAN IV
COMPUTER	UNIVAC 1108
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Mr. Sam Steinberg, Program Librarian Computer Science Center, University of Maryland, College Park, Md. 20740 Tel.: (301) 454-4261

FUNCTIONAL ABSTRACT

This system scores objective tests recorded on Digitek Optical Reader Sheets. Using an answer key it scores tests of up to 160 items for any number of students.

Output

1. A listing of students' answers to each question is optional.
2. For each student, the program lists the number of questions answered right, the number wrong, and the number omitted. Optionally included is a score which imposes a penalty for wild guessing. A standardized T score (which has an arithmetic mean of 50 and a standard deviation of 10) is also printed for each student.
3. A frequency distribution (histogram) and the mean and standard deviation of the scores are printed.
4. An item analysis is printed for *each question* on the examination.
    a) P (Difficulty Index) is the proportion of the total group who answer the question correctly. Items having a

continued

difficulty index of .50 provide maximum differentiation among students. Good tests contain items having difficulty indices in the middle range (.25 to .75).

b) D (Discrimination Index) measures the power of a single item to discriminate between the upper and lower halves of the student group, divided as to total scores. A good classroom test should have indices of discrimination of .30 or better.

Also printed are the responses to each item by the high and low scoring groups. The correct answer is indicated.

The three statistics which follow the D index are used for similar purposes and are consistent in meaning with it. All are discrimination indices expressing the relationship between item success and the total score (the criterion).

c) PHI (Phi coefficient) is an index of discriminating power when the criterion variable is a natural dichotomy and must be used as such; i.e., high-low, good-poor, etc. No assumptions are made about the form of the distribution of the group.

d) RPBI (Point biserial correlation coefficient) is an index of discriminating power when the criterion is a continuous variable. No assumptions are made about the form of the distribution of the group.

e) RB (Biserial correlation coefficient) is an index of discriminating power requiring the assumption that one of the normally distributed underlying variables has been forced into a dichotomy.

The choice among these depends partly on the purpose for which the test and item analysis data are to be used, and partly on the convenience with which each statistic serves that purpose. For most classroom examinations, the D statistic is most easily understood; it will identify items with little internal-consistency discriminating power and lead to greater efficiency of measurement when revising a test to contain the more discriminating items. It should be noted, however, that when RB and RPBI are used, the indices are usually not equal. RB tends to be substantially larger than RPBI.

continued

f) An overall measure of test reliability is provided by Kuder Richardson No. 20 and No. 21 statistics. This is an estimate of how close the same set of scores would result if the same set of items were given again. Most test writers settle with reliabilities over .60 for teacher-made tests. An index of .80 and above reveals a highly reliable examination.

Formulae

Phi Coefficient

$$\emptyset = \frac{BC - AD}{[(A + B)\ (C + D)\ (A + C)\ (B + D)]^{0.5}}$$

Point Biserial Correlation

$$r_{pbi} = \frac{\overline{X}p - \overline{X}_t}{S_t} \sqrt{\frac{p}{q}}$$

$\overline{X}$ = the mean of all scores

$\overline{X}p$ = the mean of all scores in the upper half

$\overline{X}_t$ = the mean of all scores in the lower half

p = proportion of individuals in upper half

q = proportion of individuals in lower half

S_t = standard deviation of all scores

Biserial Correlation

$$r_{bi} = \frac{\overline{X}_p - \overline{X}_t}{S_t} \times \frac{P}{y}$$

P = proportion of cases in the upper group X proportion of cases in the lower group; P = pq

y = height of ordinate of unit normal curve at point of division between p and q

Discrimination Index

$$DIS = \frac{X - Z}{N} \times 2$$

continued

X = total correct answers in upper half

Z = total correct answers in lower half

N = sample size

Kuder-Richardson 20

$$\text{K-R }(20) = \frac{NQ}{NQ-1}\,\frac{(S_t^2 - PQ)}{S_t^2}$$

Kuder Richardson 21

$$\text{K-R }(21) = \frac{NQ}{NQ-1}\left[1 - \frac{\overline{X}_t(NQ-\overline{X}_t)}{NQ\ S_t^2}\right]$$

NQ = the number of questions

PQ = p times q summed over all questions

Correction for Guessing

$$\text{Raw Score} = R - \frac{W}{A-1}$$

R = Rights

W = Wrongs

A = Number of options per item (2 to 5)

REFERENCES

Ebel, R.L., *Measuring Educational Achievements*, (Prentice-Hall, Inc., New Jersey, 1965).

Lindquist, E.F., (Ed.), *Educational Measurement*, (American Council on Education, Washington D.C., 1951).

DESCRIPTIVE TITLE	Information Retrieval Services
CALLING NAME	
INSTALLATION NAME	University of Georgia Computer Center
AUTHOR(S) AND AFFILIATION(S)	Information Science Group University of Georgia
LANGUAGE	Assembler Language & PL/1
COMPUTER	IBM 360/65
PROGRAM AVAILABILITY	Proprietary; available for use but not for distribution
CONTACT	Miss Margaret Park, Supervisory Inform. Scientist, Computer Center, The Univ. of Ga., Athens, Ga. 30601 Tel.: (404) 542-3741

FUNCTIONAL ABSTRACT

In any endeavor, scientists need to keep constantly abreast of activities in their field of interest, to be on the lookout for new ideas, and to maintain a library of useful references. The proliferation of new scientific knowledge is rapidly outpacing the capabilities of conventional information-handling and publishing techniques. Scientists are now turning to computer-based methods to help speed and channel the flow of information on a timely basis.

The Computer Center at the University of Georgia is actively engaged in establishing an Information Center. Mechanized data bases from several scientific organizations are presently available and in use at the Center. Subject areas currently represented are biology, biochemistry, nuclear science, and chemistry, including structural data files for chemical compounds. Other tape services in fields such as medicine, engineering, physics, geology, etc., will be added as interst is expressed in these subject areas.

The Computer Center's Information Sciences Unit offers assistance in creating search profiles and current awareness and retrospective searches of the scientific literature.

continued

The staff of the Computer Center includes highly trained professionals in practically every subject discipline. In addition to chemists and a microbiologist in the Information Sciences Unit, the Center also has full time staff with specialties in physics, engineering, statistics, biology, and forestry, with consulting staff available in other disciplines. All staff are also familiar with computer systems and their applications.

The Computer Center is constantly seeking better ways to satisfy the information requirements of the scientific community and we will greatly appreciate any suggestions made in this direction. New services will be added as rapidly as possible to meet newly identified needs and uses.

Persons desiring to use the Information Retrieval Services are directed to the contact person.

DESCRIPTIVE TITLE	Test-Scoring and Item-Analysis Package
CALLING NAME	(a) QUICKSCORE; (b) ITEMSTEP; (c) ITEMRS
INSTALLATION NAME	Wharton Computational Services Wharton School of Finance and Commerce University of Pennsylvania
AUTHOR(S) AND AFFILIATION(S)	Daniel Ashler Daniel Bricklin David Sheinson Wharton Computational Services
LANGUAGE	FORTRAN
COMPUTER	IBM 360/75
PROGRAM AVAILABILITY	Available to process data at Wharton Computational Services
CONTACT	Daniel Ashler, Wharton Computational Services, Wharton School of Finance and Commerce, Dietrich Hall, University of Pennsylvania, Philadelphia, Pa. 19104 Tel.: (215) 594-6422

FUNCTIONAL ABSTRACT

The Test-Scoring and Item-Analysis Package currently consists of three programs, QUICKSCORE, ITEMSTEP, and ITEMRS. Together, they provide a flexible facility for tasks that range from simple scoring of tests to the most sophisticated test analysis currently available. The best features of item-analysis programs in use on various campuses have been incorporated.

QUICKSCORE

...is the least expensive of the three programs to use. It scores tests and lists the examinees and their scores, first in alphabetical order and then in order of score. Up to 500 examinees can be scored at once. However, the input can be "batched;" that is, several groups of cards of up to 500 each can be scored in this manner, one after another. Beside each examinee's name is printed his Social Security number, his score (which has been corrected

continued

for chance success), the number of items that he got correct, that he got incorrect, that he omitted, and the number of items not reached. The corrected score gives +1 for each item answered correctly and –1/(c–1) for each answered incorrectly, where c is the number of choices per item. Omitted items are scored 0.

A brief item analysis is then performed for each question. First, the estimated fraction of the examinees who knew the right answer is computed (difficulty of question). The examinees are then partitioned into an upper and a lower group on the basis of total test score, these groups being equal or nearly equal in size. A 2X2 table is constructed as follows.

	No. Rights	No. Wrongs
Upper half	A	B
Lower half	C	D

Two indices are computed on this table: ϕ correlation coefficient (PHI) and a ratio index (RI). For a test with N choices per question and U examinees in the upper half and L examinees in the lower, the two indices are defined as follows.

$$\text{PHI} = \frac{AD - BC}{[(A+B)\ (A+C)\ (C+D)\ (B+D)]^{\frac{1}{2}}}$$

$$\text{RI} = \frac{2}{\pi} \arctan \log_2 (\text{UPSC}/\text{LOWSC}),$$

where

$$\text{UPSC} = \frac{A - B/(N-1)}{U} + \frac{1}{N-1}$$

and

$$\text{LOWSC} = \frac{C - D/(N-1)}{L} + \frac{1}{N-1}\ .$$

Both indices can have values in the range –1 to 1. The ratio index reflects the extent of discrimination between the upper and lower groups as well as the correlation of the score on the given question with total test score.

A vertical histogram of the score distribution follows, in which, however, minus scores, scores of zero, and scores of one are all counted in the score category of one. Test statistics follow, in which are given the number of questions in the test, the number of choices per item, the number of examinees, a copy of the scoring key, and the means, standard deviations, and variances of the scores.

continued

ITEMSTEP

...not only performs the same scoring as QUICKSCORE but also provides fuller item analysis and a stepwise-reduction feature. After scoring and analysis, one or more questions (items) are deleted on the basis of some criterion (supplied as a parameter), after which the shortened test is rescored and reanalyzed. The cycle of shortening and reprocessing is repeated as many times as specified. One of three different indices may be specified, on the basis of which the less desirable items shall be eliminated. The first is the W index, which is a weighted combination of the Davis[1] difficulty and discrimination indices. The second is the P-adjusted index, which is the proportion of correct responses, adjusted or corrected for chance success. The third is the contribution of the item to the total test variance. The user specifies either the index level required for retention of an item or that some fixed number of items with the worst index values be dropped at each cycle. A reliability coefficient is also computed. The stepwise-reduction process tends to produce successive score distributions that are more and more rectangular.

ITEMRS

...is the most advanced item-analysis program presently available. First, it provides a header page, in which the indices are explained. Then there are optional printouts of the main matrices produced during the run: namely, the sum-of-products matrix, variance—covariance matrix, and the product—moment correlation matrix. (It should be noted that the correlations are product—moment and not point biserial because the items are not scored dichotomously.) Next, an optional cluster analysis is performed on the correlations after they have been transformed to Fisher's Z's. (A cluster analysis is quicker and less expensive than factor analysis and for many purposes is just about as useful.) Each pass produces larger clusters and fewer of them.

Next, the examinee's scores are listed in a tabulation like that of QUICKSCORE, in order by name and then by score. Alongside the listing of the scores are 16 blank columns. If desired, data may be printed in these columns, such as scaled scores, grades, or comments. (For example, to print scaled scores, one provides a card for each possible score that can be obtained, specifying what shall be printed for that score.)

A large number of indices are computed for each item, including P (proportion of correct responses) and Q (proportion of incorrect responses), both unadjusted and adjusted for chance success; the variance contribution of the item; the product—moment correlation of the item with the total score; the "latent" correlation, which

continued

is the estimate of the correlation between the underlying ability to answer the item correctly and the total score; and the Fisher's Z transform of the product–moment correlation. Also, the effective number of choices (that is, the number of choices that actually attracted examinees) the estimated proportion who knew the answer (based on the effective number of choices), the Davis difficulty and discrimination indices, the W index, which is a weighted combination of the Davis indices, and the index DICAP, which measures the deviation of the responses to this item from the ideal pattern in which 50% of the examinees would choose the correct alternative and the remaining 50% would distribute equally over the remaining alternatives. Optionally, a tabulation of the number of responses to each alternative is given for each fifth of the examinees.

Following the item analysis, a histogram is printed, followed by test statistics, including the covariance reliability coefficient, the KR-20 reliability coefficient, the standard error of measurement, the variance error of measurement, and the test mean, standard deviation, and variance of total test scores.

REFERENCES

1. Davis, F.B., *Item Analysis Data* (Harvard University Press, Cambridge, Mass., 1949).

DESCRIPTIVE TITLE	Cross-Cultural Comparison
CALLING NAME	POLYCOMP
INSTALLATION NAME	Dartmouth College Kiewit Computation Center
AUTHOR(S) AND AFFILIATION(S)	Written by William Koenig, Class of 1970, for James Fernandez, Department of Anthropology, Dartmouth College
LANGUAGE	Dartmouth BASIC
COMPUTER	GE–635
PROGRAM AVAILABILITY	Magnetic tape and listings presently available
CONTACT	A. Kent Morton, EIN Technical Representative, Kiewit Computation Center, Dartmouth College, Hanover, N.H. 03755 Tel.: (603) 646-2864

FUNCTIONAL ABSTRACT

POLYCOMP is designed for cross-cultural comparison using the data contained in Murdock's *Ethnographic Atlas*.[1] (If a copy of the "Ethnographic Atlas Codesheet"[2] is not immediately available, selected portions of it may be examined during the running of the program, or by running ETH–CODE, EIN No. 000 0119.) The user has the option of working with either one or two groups of cultures (or political units). Up to 20 cultures may be assigned to each group. A comparison is first made for each unit with each of the 92 characteristics from the *Ethnographic Atlas* compared for each culture against every other culture in that group. Characteristics with "insufficient information" are discounted.

The user then has the option of having the identity matrices printed out in one of two formats, after which he may weight the characteristics according to his own preference and obtain a new comparison table.

POLYCOMP may be used in conjunction with ETH–CODE, ETH–DGRE, ETH–RAND, ETH–INFO, ETHATLAS, CULTCOMP, and CULTPIK (EIN Nos. 000 0119–000 0125).

continued

REFERENCES

1. Murdock, G.P., *Ethnology Atlas*, (Univ. of Pittsburgh Press, Pittsburgh, Pa., 1967).

2. "Ethnographic Atlas Codesheet", (Depart. of Anthrop., College Museum, Dartmouth College, Hanover, N.H., 1970), Revised. Available from the EIN Office at the cost of reproduction and mailing.

3. Murdock, G.P., "Ethnographic Atlas: A Summary", *Ethnology, VI*, 2, (April 1967).

4. "Ethnographic Atlas", by the editors of *Ethnology, I*, 1, (Jan., 1962)–*VII*, 3, (July, 1968), installments of information in each issue.

5. Barry III, H., "Ethno", (*Ethnology* version of *Ethnographic Atlas*, Univ. of Pittsburgh, Pittsburgh, Pa., 1970).

DESCRIPTIVE TITLE	Major characteristics from Murdock's *Ethnographic Atlas*
CALLING NAME	ETH–CODE
INSTALLATION NAME	Dartmouth College Kiewit Computation Center
AUTHOR(S) AND AFFILIATION(S)	Program written by William Koenig, Class of 1970, for James Fernandez, Department of Anthropology, Dartmouth College Codesheet revised from that of Herbert Barry III, Department of Anthropology, University of Pittsburgh
LANGUAGE	Dartmouth BASIC
COMPUTER	GE–635
PROGRAM AVAILABILITY	Magnetic tape and listings presently available
CONTACT	A. Kent Morton, EIN Technical Representative, Kiewit Computation Center, Dartmouth College, Hanover, N.H. 03755 Tel.: (603) 646-2864

FUNCTIONAL ABSTRACT

ETH–CODE is a driver program for the sub-program CODESHEET, which contains the 48 major characteristics from the *Ethnographic Atlas*[1-5]. After these have been listed, the user has the option to see the individual codings and subdivisions for any of these major characteristics. The output takes the form of a list with matching column and code numbers as they appear in "Ethnographic Atlas Codesheet"[2]. This program is, in effect, a partial replacement for the complete codesheet, if the user does not have one available. Requests to have individual codings output should be limited to only a few of the major characteristics, as the complete codesheet produces many pages of output and requires considerable time.

REFERENCES

1. Murdock, G.P., *Ethnology Atlas*, (Univ. of Pittsburgh Press, Pittsburgh, Pa., 1967).

continued

2. "Ethnographic Atlas Codesheet", (Depart. of Anthrop., College Museum, Dartmouth College, Hanover, N.H., 1970), Revised. Available from the EIN Office at the cost of reproduction and mailing.

3. Barry III, H., "Ethno", (*Ethnology* version of *Ethnographic Atlas*, Univ. of Pittsburgh, Pittsburgh, Pa., 1970).

4. "Ethnographic Atlas", by the editors of *Ethnology, I*, 1(Jan., 1962)-*VII*, 3,(July, 1968), installments of information in each issue.

5. Murdock, G.P., "Ethnographic Atlas: A Summary", *Ethnology, VI*, 2,(April, 1967).

DESCRIPTIVE TITLE	Random Samples of Cultures
CALLING NAME	ETH–DGRE
INSTALLATION NAME	Dartmouth College Kiewit Computation Center
AUTHOR(S) AND AFFILIATION(S)	Program written by William Koenig, Class of 1970, for James Fernandez, Department of Anthropology, Dartmouth College Data collected by Daniel Gordon, Department of Sociology, University of Oregon, from Murdock's *Ethnographic Atlas*
LANGUAGE	Dartmouth BASIC
COMPUTER	GE-635
PROGRAM AVAILABILITY	Magnetic tape and listings presently available
CONTACT	A. Kent Morton, EIN Technical Representative, Kiewit Computation Center, Dartmouth College, Hanover, N.H. 03755 Tel.: (603) 646-2864

FUNCTIONAL ABSTRACT

One of the chief problems in making cultural comparisons and gathering a random sample of cultures is making sure they are not so close together as to share many characteristics due to diffusion and common environmental pressures. In order to avoid this, cultural anthropologists employ a 3–degree rule, to insure a geographical separation. Respecting this rule, ETH–DGRE randomly selects one culture (if any cultures have been recorded[1] to exist in the sample) in each of the 72 cells of 30–degree latitude and longitude on mercator projection. Each selected culture is then compared with the cultures of the neighboring cells to check for the 3–degree rule. This includes checking cultures of the far west with cultures of the far east, and those of the far north with those of the far south, when such cultures exist.

Output is in the form of a simple list or a two-page map with 30–degree cells indicated.

continued

The user cannot be guaranteed to get a sample of given size by using ETH—DGRE, whereas ETH—RAND (EIN No. 000 0121) will supply a sample of 50 cultures (and consume considerably more time in doing so).

Other programs which might be useful in conjunction with this one include ETH—INFO, ETH—CODE, ETHATLAS, CULTCOMP, CULTPIK, and POLYCOMP, (EIN No. 000 0118 - 000 0119, 000 0122 - 000 0125).

REFERENCES

1. "Condensed Geographic Codesheet of the Ethnographic Atlas: 861 Cultures", (Depart. of Anthrop., College Museum, Dartmouth College, Hanover, N.H., 1970).

2. McNett, C.W., and Kirk, R.E., "Drawing Random Samples in Cross-Cultural Studies; A Suggested Method", Amer. Anthrop., *70*, 1, (Feb. 1968).

DESCRIPTIVE TITLE	Random Sample of 50 Cultures
CALLING NAME	ETH-RAND
INSTALLATION NAME	Dartmouth College Kiewit Computation Center
AUTHOR(S) AND AFFILIATION(S)	Program written by William Koenig, Class of 1970, for James Fernandez, Department of Anthropology, Dartmouth College Data collected by Daniel Gordon, Department of Sociology, University of Oregon, from Murdock's *Ethnographic Atlas*
LANGUAGE	Dartmouth BASIC
COMPUTER	GE-635
PROGRAM AVAILABILITY	Magnetic tape and listings presently available
CONTACT	A. Kent Morton, EIN Technical Representative, Kiewit Computation Center, Dartmouth College, Hanover, N.H. 03755 Tel.: (603) 646-2864

FUNCTIONAL ABSTRACT

ETH-RAND provides an alternative to ETH-DGRE (EIN No. 000 0120) in randomly selecting a sample of 50 cultures separated by the 3-degree rule. The program initially selects a random point of specific latitude and longitude. With this point as an origin the program then lays out a mercator projection and selects 50 points by randomly choosing latitude and longitude points which are then converted to points on the normal mercator projection map. ETH-RAND then examines the 30-degree cell in which this point lies and the eight surrounding cells to determine the culture closest to this point. This culture must pass the 3-degree rule or the program persists until it has found the closest culture which does pass the test. The program then prints out the coordinates of the random point, followed by the selected culture and its coordinates.

continued

The program CULTCOMP (EIN No. 000 0124) and ETH-INFO (EIN No. 000 0122) may be employed to gather additional information on these cultures.

REFERENCES

"Condensed Geographic Code Sheet of the Ethnographic Atlas: 861 Cultures", (Depart. of Anthrop., College Museum, Dartmouth College, Hanover, N.H., 1970).

McNett, C.W., and Kirk, R.E., "Drawing Random Samples in Cross-Cultural Studies; A Suggested Method", Amer. Anthrop., *70*, 1, (Feb., 1968).

DESCRIPTIVE TITLE	Demographic Information from Murdock's *Ethnographic Atlas*
CALLING NAME	ETH–INFO
INSTALLATION NAME	Kiewit Computation Center Dartmouth College
AUTHOR(S) AND AFFILIATION(S)	Program written by William Koenig, Class of 1970, for James Fernandez, Department of Anthropology, Dartmouth College Data collected by Daniel Gordon, Department of Sociology, University of Oregon, from Murdock's *Ethnographic Atlas*
LANGUAGE	Dartmouth BASIC
COMPUTER	GE–635
PROGRAM AVAILABILITY	Magnetic tape and listings presently available
CONTACT	A. Kent Morton, EIN Technical Representative, Kiewit Computation Center, Dartmouth College, Hanover, N.H. 03755 Tel.: (603) 646-2864

FUNCTIONAL ABSTRACT

ETH-INFO provides information on any of 861 cultures contained in the random access file DEGRE. Output consists of a table which shows the culture(s) of interest to the user, regional identification, latitude, longitude, size of population, year data were collected, and the year the population estimate was made. Definitions of regional identifications are output on request.

Other programs which might be useful in conjunction with ETH–INFO include ETH–CODE, ETH–DGRE, ETH–RAND, ETHATLAS, CULTCOMP, CULTPIK, and POLYCOMP, (EIN Nos. 000 0118 - 000 0125).

REFERENCES

Murdock, G.P., *Ethnology Atlas*, (Univ. of Pittsburgh Press, Pittsburgh, Pa., 1967).

continued

Barry III, H., "Ethno", (*Ethnology* version of *Ethnographic Atlas*, Univ. of Pittsburgh, Pittsburgh, Pa., 1970).

"Ethnographic Atlas", by the editors of *Ethnology*, *I*, 1,(Jan., 1962)–*VII*, 3, (July, 1968), installments of information in each issue.

Murdock, G.P., "Ethnographic Atlas: A Summary", *Ethnology*, *VI*, 2, (April, 1967).

DESCRIPTIVE TITLE	Comparison of Characteristics from "Ethnographic Atlas"
CALLING NAME	ETHATLAS
INSTALLATION NAME	Dartmouth College Kiewit Computation Center
AUTHOR(S) AND AFFILIATION(S)	James Fernandez, Department of Anthropology, Dartmouth College Modified by Mark Hebenstreit, Class of 1970 Data coded by Herbert Barry III, Department of Anthropology, University of Pittsburgh
LANGUAGE	Dartmouth BASIC
COMPUTER	GE—635
PROGRAM AVAILABILITY	Magnetic tape and listings presently available
CONTACT	A. Kent Morton, EIN Technical Representative, Kiewit Computation Center, Dartmouth College, Hanover, N.H. 03755 Tel.: (603) 646-2864

FUNCTIONAL ABSTRACT

ETHATLAS scans data from the "Ethnographic Atlas" (92 characteristics of 1168 societies). It will compare any two characteristics on a presence or absence basis and compute degrees and significance of association. Any five conditions can be held constant. Thus, for example, the association may be computed for only those societies with intense agriculture or with patrilineal descent.

The program may be run only in consultation with the "Ethnographic Atlas Codesheet"[1] obtainable from the Department of Anthropology. If a codesheet is not immediately available, selected portions of it may be examined by running ETH—CODE, (EIN No. 000 0119). The codesheet gives the column numbers of the various characteristics and the coding relevant to decisions as to their presence or absence.

continued

REFERENCES

1. "Ethnographic Atlas Codesheet", (Depart. of Anthrop., College Museum, Dartmouth College, Hanover, N.H., 1970), Revised. Available from the EIN Office at the cost of reproduction and mailing.

2. Murdock, G.P., *Ethnology Atlas*, (Univ. of Pittsburgh Press, Pittsburgh, Pa., 1967).

3. Barry III, H., "Ethno", (*Ethnology* version of *Ethnographic Atlas*, Univ. of Pittsburgh, Pittsburgh, Pa., 1970).

4. "Ethnographic Atlas", by the editors of *Ethnology, I*, 1, (Jan., 1962)-*VII*, 3, (July, 1968), installments of information in each issue.

5. Murdock, G.P., "Ethnographic Atlas: A Summary", *Ethnology, VI*, 2, (April, 1967).

DESCRIPTIVE TITLE	Cultural Comparison
CALLING NAME	CULTCOMP
INSTALLATION NAME	Dartmouth College Kiewit Computation Center
AUTHOR(S) AND AFFILIATION(S)	Program written by William Koenig, Class of 1970, for James Fernandez, Department of Anthropology, Dartmouth College Revised by Mark Hebenstreit, Class of 1970 Data from the "Ethnographic Atlas" were obtained from Herbert Barry III, Department of Anthropology, University of Pittsburgh
LANGUAGE	Dartmouth BASIC
COMPUTER	GE-635
PROGRAM AVAILABILITY	Magnetic tape and listings presently available
CONTACT	A. Kent Morton, EIN Technical Representative, Kiewit Computation Center, Dartmouth College, Hanover, N.H. 03755 Tel.: (603) 646-2864

FUNCTIONAL ABSTRACT

CULTCOMP retrieves coded data from George Peter Murdock's *Ethnographic Atlas* and translates them for cultural comparison. The user learns the degree of similarity and dissimilarity between the two cultures of his choice, and has the option to see a print-out in chart form which gives relevant characteristics for each culture. Data are available on 1168 cultures. In the event a codesheet is not available, the user has the option to list all 1168 cultures for examination.

REFERENCES

Murdock, G.P., *Ethnology Atlas*, (Univ. of Pittsburgh Press, Pittsburgh, Pa., 1967).

continued

Barry III, H., "Ethno", (*Ethnology* version of *Ethnographic Atlas*, Univ. of Pittsburgh, Pittsburgh, Pa., 1970).

"Ethnographic Atlas", by the editors of *Ethnology, I*, 1,(Jan., 1962)–*VII*, 3, (July, 1968), installments of information in each issue.

Murdock, G.P., "Ethnographic Atlas: A Summary", *Ethnology, VI*, 2,(April, 1967).

DESCRIPTIVE TITLE	Culture Identification
CALLING NAME	CULTPIK
INSTALLATION NAME	Dartmouth College Kiewit Computation Center
AUTHOR(S) AND AFFILIATION(S)	Program written by William Koenig, Class of 1970, for James Fernandez, Department of Anthropology, Dartmouth College
LANGUAGE	Dartmouth BASIC
COMPUTER	GE-635
PROGRAM AVAILABILITY	Magnetic tape and listings presently available
CONTACT	A. Kent Morton, EIN Technical Representative, Kiewit Computation Center, Dartmouth College, Hanover, N.H. 03755 Tel.: (603) 646-2864

FUNCTIONAL ABSTRACT

CULTPIK allows the user to identify those cultures which satisfy any selected set of characteristics. Characteristics may be selected from Murdock's *Ethnographic Atlas*[1] or from Textor's *Cross-Cultural Summary*[2]. Complete codesheets may be obtained from the Department of Anthropology, College Museum, Dartmouth College. Portions of the codesheets may be examined by running either ETH—CODE (EIN No. 000 0119) or TEX—CODE (EIN No. 000 0127).

REFERENCES

1. Murdock, G.P., *Ethnographic Atlas*, (Univ. of Pittsburgh Press, Pittsburgh, Pa., 1967).

2. Textor, R.B., Comp., *A Cross Cultural Summary*, (HRAF Press, New Haven, Conn., 1967).

3. "Cross-Cultural Survey", Codesheet for R.B. Textor's finished characteristics from *A Cross-Cultural Summary*, (Dept. of Anthrop., College Museum, Dartmouth College, Hanover, N.H.). Available from the EIN Office at the cost of reproduction and mailing.

continued

4. Barry III, H., "Ethno", (*Ethnology* version of *Ethnographic Atlas*, Univ. of Pittsburgh, Pittsburgh, Pa., 1970).

5. "Ethnographic Atlas" by the editors of *Ethnology, I*, 1,(Jan., 1962)—*VII*, 3,(July, 1968), installments of information in each issue.

6. Murdock, G.P., "Ethnographic Atlas: A Summary", *Ethnology, VI*, 2, (April, 1967).

DESCRIPTIVE TITLE	Finished Characteristics of 400 Cultures
CALLING NAME	TEXTOR
INSTALLATION NAME	Dartmouth College Kiewit Computation Center
AUTHOR(S) AND AFFILIATION(S)	Written by William Koenig, Class of 1970, for James Fernandez, Department of Anthropology, Dartmouth College Data for the 400 cultures were compiled by Robert B. Textor
LANGUAGE	Dartmouth BASIC
COMPUTER	GE–635
PROGRAM AVAILABILITY	Magnetic tape and listings presently available
CONTACT	A. Kent Morton, EIN Technical Representative, Kiewit Computation Center, Dartmouth College, Hanover, N.H. 03755 Tel.: (603) 646-2864

FUNCTIONAL ABSTRACT

TEXTOR reproduces the dichotomies for the finished characteristics of 400 cultures as constructed by Robert B. Textor[1]. The user has the option of seeing a list of the major characteristic categories, as well as selected sub-divisions, just as in TEX–CODE (EIN No. 000 0127).

REFERENCES

1. Textor, R.B., Comp., *A Cross-Cultural Summary*, (HRAF Press, New Haven, Conn., 1967).

DESCRIPTIVE TITLE	Finished characteristics from Textor's *A Cross-Cultural Summary*
CALLING NAME	TEX-CODE
INSTALLATION NAME	Dartmouth College Kiewit Computation Center
AUTHOR(S) AND AFFILIATION(S)	Written by William Koenig, Class of 1970, for James Fernandez, Department of Anthropology, Dartmouth College
LANGUAGE	Dartmouth BASIC
COMPUTER	GE—635
PROGRAM AVAILABILITY	Magnetic tape and listings presently available
CONTACT	A. Kent Morton, EIN Technical Representative, Kiewit Computation Center, Dartmouth College, Hanover, N.H. 03755 Tel.: (603) 646—2864

FUNCTIONAL ABSTRACT

Forty-four major divisions of Textor's finished characteristics[1] are first printed out, after which the user is given the chance to see selected portions in detail.

TEX—CODE may be used in conjunction with TEXTOR, (EIN No. 000 0126), or CULTPIK, (EIN No. 000 0125).

REFERENCES

1. Textor, R.B., Comp., *A Cross-Cultural Summary*, (HRAF Press, New Haven, Conn., 1967).

DESCRIPTIVE TITLE	Biomedical Computer Programs
CALLING NAME	BMD
INSTALLATION NAME	Washington University Computing Facilities
AUTHOR(S) AND AFFILIATION(S)	Health Sciences Computing Facility University of California, Los Angeles
LANGUAGE	FORTRAN IV
COMPUTER	IBM 360/50
PROGRAM AVAILABILITY	Proprietary; available for use at Washington University. Available for distribution from authors.
CONTACT	Dr. C. B. Drebes, Mgr., Scientific Data Processing, Computing Facilities, Box 1098, Washington University, St. Louis, Mo. 63130 Tel.: (314) 863-0100 ext. 3141

FUNCTIONAL ABSTRACT

The BMD package contains a variety of statistical programs in the following areas.

Description and Tabulation

Multivariate Analysis
a. Factor Analysis
b. Discriminant Analysis
c. Canonical Analysis

Regression Analysis
a. Linear
b. Polynomial
c. Asymptotic

Analysis of Variance and Covariance

Time Series Analysis

Special Programs
a. Life Table and Survival Rate
b. Contingency Table Analysis
c. Biological Assay
d. Guttman Scaling

continued

REFERENCES

Dixon, W.J., Ed., *BMD: Biomedical Computer Programs* (Berkeley: Univ. of Cal. Press, 1968).

Dixon, W.J., Ed., *BMD: Biomedical Computer Programs, X-series Supplement* (Berkeley: Univ. of Cal. Press, 1969).

DESCRIPTIVE TITLE	General Purpose Simulation System
CALLING NAME	GPSS
INSTALLATION NAME	Washington University Computing Facilities
AUTHOR(S) AND AFFILIATION(S)	IBM Application Program
LANGUAGE	360 Assembly Language
COMPUTER	IBM 360/50
PROGRAM AVAILABILITY	Proprietary; available for use at Washington University. Available for distribution from IBM.
CONTACT	Dr. C.B. Drebes, Mgr., Scientific Data Processing, Computing Facilities, Box 1098, Washington University, St. Louis, Mo. 63130 Tel.: (314) 863-0100 ext. 3141

FUNCTIONAL ABSTRACT

GPSS is a transaction-oriented language designed for conducting evaluations and experiments concerning the behavior of systems, methods and processes. It has a modular structure which permits "transactions" to flow through the system, where their interactions can be observed and modified. A "clock" is maintained by which events are either scheduled to occur or else determined by one of the eight random number generators provided. Information can be obtained regarding sequencing of operations, scheduling and allocation rules, inventories, queuing disciplines, machine failures, etc. In general, various trade-offs between cost and performance can be studied.

REFERENCES

General Purpose Simulation System/360, Introductory User's Manual (IBM Manual GH20-0304-0), (White Plains, N.Y.: IBM Corp. Tech. Pub. Depart., 1969).

General Purpose Simulation System/360, User's Manual (IBM Manual GH20-0326-3), (White Plains, N.Y.: IBM Corp. Tech. Pub. Depart., 1969).

General Purpose Simulation System/360, OS (360A-CS-17X) Operator's Manual (IBM Manual H20-0311), (White Plains, N.Y.: IBM Corp. Tech. Pub. Depart., 1969).

DESCRIPTIVE TITLE	A Primal-Dual Transportation Algorithm
CALLING NAME	TRANSPRT
INSTALLATION NAME	Washington University Computing Facilities
AUTHOR(S) AND AFFILIATION(S)	D.E. Burlingame Washington University Computing Facilities
LANGUAGE	FORTRAN IV
COMPUTER	IBM 360/50
PROGRAM AVAILABILITY	Deck and listing presently available
CONTACT	Dr. C.B. Drebes, Mgr., Scientific Data Processing, Computing Facilities, Box 1098, Washington University, St. Louis, Mo. 63130 Tel.: (314) 863-0100 ext. 3141

FUNCTIONAL ABSTRACT

TRANSPRT solves a standard transportation problem using the primal-dual transportation algorithm. This program solves a model in which the objective is to "transport" a single commodity from various origins to different destinations at a minimum total shipping cost. The availability at each origin, the demand at each destination, and the cost to ship one unit of the project from any origin to any destination are required inputs to this model. This model can be applied to certain other types of industrial or business problems that have nothing to do with shipping. Personnel assignment, machine assignment, product and inventory scheduling are a few such applications.

The program utilizes integer arithmetic and requires integer input. If the sum of the origin availabilities is not equal to the sum of the destination requirements, an artificial origin or destination is set up, with zero costs, to handle the excess. The program will handle up to 50 origins and 150 destinations, using 94K bytes of core storage.

REFERENCES

Hadley, G., *Linear Programming* (Reading, Mass: Addison-Wesley Pub. Co., Inc., 1962).

DESCRIPTIVE TITLE	FORTRAN Program for Computer Based Serials Holdings Management
CALLING NAMES	UPDATE, HOLD, PUBLSH, LANSUB
INSTALLATION NAME	Indiana University-Purdue University at Indianapolis Research Computation Center
AUTHOR(S) AND AFFILIATION(S)	Mrs. Alma Connell School of Medicine Library Indiana University Mrs. J. Mueller School of Medicine Library Indiana University Mrs. Judy Silence Research Computation Center, IUPUI
LANGUAGE	IBM FORTRAN IV (IBSYS)
COMPUTER	IBM 7040
PROGRAM AVAILABILITY	Deck and listing presently available
CONTACT	Dr. David A. Neal, EIN Tech. Rep., Research Comp. Ctr., Indiana Univ.—Purdue Univ. at Indianapolis, 1100 West Michigan St., Indianapolis, Ind. 46202 Tel.: (317) 639-7813

FUNCTIONAL ABSTRACT

This system is a long range project concerning the holdings of serials in the School of Medicine Library. Three overall applications are (1) updating the file to maintain a current holding list, (2) publication of selected lists using search techniques, and (3) annual publication of complete holdings list.

Several programs have been prepared for these applications. Detailed descriptions are given in Ref. 1. A general description of each program follows.

System Update Program 1

This program produces a readable list for visual checking of every item on the cards. When a set of cards has been keypunched,

continued

they are run with the System Update Program 1 using the card reader as the input unit. The output is checked and the cards are corrected. The cards are then added to the master file with the UPDATE (number 2) program.

System Update Program 2

This program builds a master tape which may be used as data for the programs that follow. UPDATE will also be used for maintaining the current file, i.e., inserting new cards as they are punched and replacing cards when new information is needed.

HOLD

HOLD can be run with the master file to produce a listing of the information on each card on that file. HOLD also provides a summary table giving the number of titles for specified categories.

PUBLSH

PUBLSH provides thirteen options for listing holdings for publication. These are (1) full list with holdings, (2 and 3) nursing list with or without holdings, (4 and 5) currently received list with or without holdings, (6 and 7) Index Medicus list with or without holdings, (8 and 9) Indexes, Abstracts and Bibliographies list with or without holdings, (10 and 11) International Nursing Index with or without holdings, and (12 and 13) microform with or without holdings. Multiple copies of output from this program are available.

LANSUB

LANSUB provides four general options for title listings. These are (1) language only, (2) one subject, (3) two subjects listed according to "and" or "or" logic, and (4) a language and a subject cross-indexed.

Listings of complete holdings or of currently received serials only can be obtained under the four options above. On complete holdings lists, currently received titles are indicated by a plus (+) to the left of the title. Multiple copies of output from this program are available.

REFERENCE

1. Indiana University Medical Center Research Computation Center Library, *Program Description* (Indianapolis, Ind: Indiana Univ., 1966, Rev. 1968). Available from the EIN Office at the cost of reproduction and mailing.

Computer Utility

DESCRIPTIVE TITLE	Upper- and Lower-Case Greek and Roman Alphabetic Plotting
CALLING NAME	SCRIPT
INSTALLATION NAME	University of Notre Dame Computing Center
AUTHOR(S) AND AFFILIATION(S)	Nicholas DiCianni
LANGUAGE	FORTRAN
COMPUTER	UNIVAC 1107
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Leonard Kilian Elizabeth Hutcheson Computing Center University of Notre Dame

FUNCTIONAL ABSTRACT

The SCRIPT subroutine provides useful and flexible character sets for the CalComp plotter. Two fonts are available, Roman and Greek, each in upper and lower case. Special Control Characters allow the user to change case or font and permit backspacing, superscripting, and subscripting. The routine also provides for plotting at angles. Italics may be specified also.

DESCRIPTIVE TITLE	Program Deck Identification Field Sequencing Program
CALLING NAME	IDSEQN NUCC128
INSTALLATION NAME	Vogelback Computing Center, Northwestern University
AUTHOR(S) AND AFFILIATION(S)	Vogelback Computing Center, Northwestern University
LANGUAGE	CDC FORTRAN IV
COMPUTER	CDC 6400
PROGRAM AVAILABILITY	Deck and listing presently available
CONTACT	Lorraine Borman, EIN Technical Representative, Vogelback Computing Center, Northwestern University, 2129 Sheridan Road, Evanston, Ill. 60201. Tel.: (312) 492-3682

FUNCTIONAL ABSTRACT

The program reproduces FORTRAN, COMPASS, and ASCENT source decks, with the ID field (Cols. 73-80) containing

(1) either a 2-, 3-, or 4-alphanumeric-character identification code;
(2) either a 1- or 2-digit subprogram number (this may be omitted if desired);
(3) either a 3- or 4-digit card-sequence number within each program.

Each time an END card (END may be in any Cols. 7-72) is encountered, the subprogram number is incremented by one. Multiple decks (with a separator card and a new control card between each deck) can be run at one time. Printed output consists of a listing of the sequenced deck (each subprogram beginning on a new page), with the original ID field printed optionally to the right of each card listing.

DESCRIPTIVE TITLE	Sequence Checking
CALLING NAME	SEQCHK NUCC078
INSTALLATION NAME	Vogelback Computing Center, Northwestern University
AUTHOR(S) AND AFFILIATION(S)	Betty Benson, Vogelback Computing Center, Northwestern University
LANGUAGE	CDC FORTRAN IV
COMPUTER	CDC 6400
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Lorraine Borman, EIN Technical Representative, Vogelback Computing Center, Northwestern University, 2129 Sheridan Road Evanston, Ill. 60201 Tel.: (312) 492-3682

FUNCTIONAL ABSTRACT

Inasmuch as the data for each of a number of subjects or cases in an analysis may occupy several cards, it is clearly necessary that these data cards be read in the same order for each case. SEQCHK checks to see if a data deck (or a data tape) contains the correct number of cards, in the correct order, for each case, up to a maximum of 9999 cases. An appropriate error message is printed whenever a card is found out of order or if any cards are missing.

DESCRIPTIVE TITLE	Test Scorer and Statistical Analysis
CALLING NAME	TSSA NUCC072
INSTALLATION NAME	Vogelback Computing Center Northwestern University
AUTHOR(S) AND AFFILIATION(S)	Richard Wolf and Leopold Klopfer The University of Chicago K. Jones Harvard University Graduate School of Education A. Gasche, B. Wright, and C. Bradford The University of Chicago
LANGUAGE	CDC FORTRAN IV
COMPUTER	CDC 6400
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Lorraine Borman, EIN Technical Representative, Vogelback Computing Center, Northwestern University, 2129 Sheridan Road, Evanston, Ill. 60201 Tel.: (312) 492-3682

FUNCTIONAL ABSTRACT

TSSA scores multiple-response tests that have a single correct response for each item, computes test and item statistics, computes the tetrachoric interitem correlation matrix, and performs a factor analysis and varimax rotation. The point-biserial correlation of each item with a criterion score also may be obtained. The number of alternative responses may vary from item to item.

Multiple-scoring keys may be used with the same set of data cards. This feature makes TSSA applicable for scoring and analysis of the Kuder preference record and similarly constructed multikeyed instruments. In addition, multiple jobs may be run.

Computational results include the following.

Individual Scores, including raw scores and scores corrected for guessing

continued

Item-Response Information, including the proportion of subjects selecting the correct response for each item (i.e., difficulty) and the proportion of subjects selecting each response

Test Statistics,[1] including the mean, standard deviation, skewness, and kurtosis (and their standard errors computed from formulas outlined by Fisher); test reliability (computed by Kuder–Richardson formula 20); a validity coefficient; and a Pearson product–moment correlation

Item-Analysis Information,[2,3] including the proportion of subjects passing an item, the item standard deviation, point-biserial correlations, reliability index, and validity index (computed from formulas given by Guilford and Gullicksen)

Tetrachoric Interitem Correlations[4]

Factor Analysis of Interitem Correlation Matrix,[5] with varimax rotation and plot

Multiple-Scoring Keys

A special feature of TSSA is the use of Multiple-Scoring Keys. Such a feature has two uses. (1) It allows a single set of responses to be scored in more than one way; practical application of this is the use of the program with data from an instrument such as the Kuder preference record where several scores, based on analyses of the same set of items, are desired. (2) A more common situation is the use of the program to obtain subtest scores and a total test score from one instrument. An application of this might be to score all the responses on a reading test and to obtain scores for reading speed, reading vocabulary, and level of comprehension, as well as a total reading score.

Multiple Jobs

The limitation of the procedures in the analyses of subtests is that only the raw scores for each individual or each subject will be meaningful. Corrected scores will *not* have any meaning. Thus, if an analysis excludes any items initially read, no corrected score will be printed. The reason for this is that the program treats *all* items excluded from analysis as incorrect responses. Thus, the use of a formula for correcting scores for guessing is fallacious in this instance. If the user desires meaningful corrected scores for subtests, he will have to reproduce his Data Cards and run multiple jobs, selecting for reading on any one job only those items for which analysis is desired, i.e., the subtest.

continued

The use of the Multiple-Jobs feature for scoring subtests yields meaningful corrected scores, as already mentioned. However, the advantage of the Multiple-Scoring-Key feature of the program is that the data need be read from cards only once. Subsequent reading of the data is done from binary tape (tape B3), thus achieving savings of time and, hence, cost.

The Multiple-Jobs feature may also be used for scoring and analyzing severly entirely different tests, with only a single loading of the program.

REFERENCES

1. Fisher, R.A., *Statistical Methods for Research Workers* (Oliver & Boyd, Edinburgh, 1954), 12th ed., pp. 70–75.

2. Guilford, J.P., *Psychometric Methods* (McGraw–Hill Book Co., Inc., New York, 1954), pp. 373–464.

3. Gullicksen, H., *Theory of Mental Tests* (John Wiley & Sons, Inc., New York, 1950), pp. 363–396.

4. Farrell, R.H., and Stern, G.G., "A Measure of Tetrachoric Association," presented at APA convention, Sept. 1951, Chicago (unpublished; a copy of this paper may be obtained by writing The University of Chicago Computing Center).

5. Wright, B.D., "120X120 Principle Components Analysis: Symmetric (with Varimax Rotation and Plot)," Univ. Chicago Soc. Sci. Div. Computer Library (1963; unpublished).

DESCRIPTIVE TITLE	Simulator of SAMOS (A Simple Imaginary Machine Language for Instruction)
CALLING NAME	SAMOS
INSTALLATION NAME	Computing Center, The Florida State University.
AUTHOR(S) AND AFFILIATION(S)	Marilyn Zupsich, Computing Center, The Florida State University
LANGUAGE	FORTRAN IV
COMPUTER	CDC 6400, 65K
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Marilyn Zupsich, Computing Center, The Florida State University, Tallahassee, Fla. 32306 Tel.: (904) 599-3418

FUNCTIONAL ABSTRACT

This program simulates the SAMOS computer described in the textbooks referred to below. The SAMOS programming language is a simplified machine-type language and is used to illustrate the logic of a computer to beginning-programming students. The simulator accepts programs written in the SAMOS language and executes them just as the mythical SAMOS computer would.

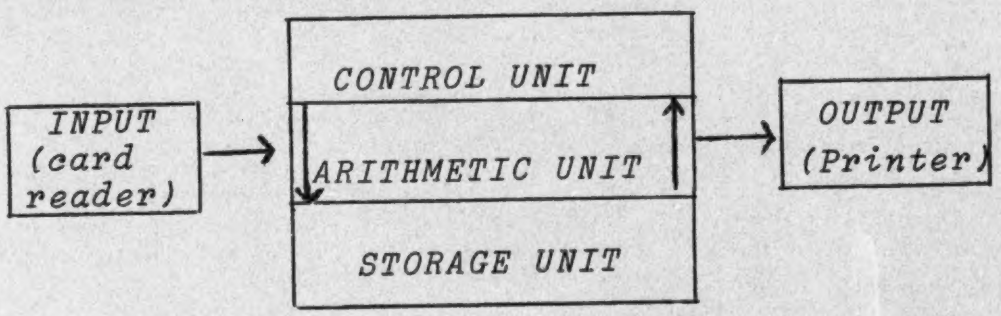
The SAMOS computer has 10,000 words of magnetic-core storage, each of which contains 10 characters and a sign (+ or -). Fifteen instruction types, covering arithmetic, branching, input, output, shifting, and the use of index registers, are simulated.

The basic design of the SAMOS computer is described in the following diagrams.

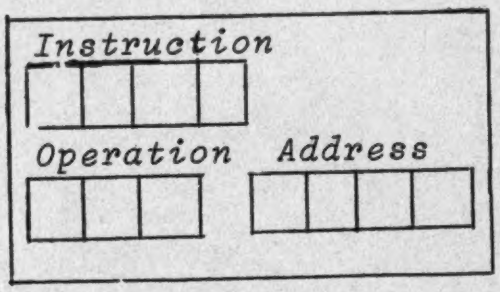
continued

Description of SAMOS Computer

Basic Design

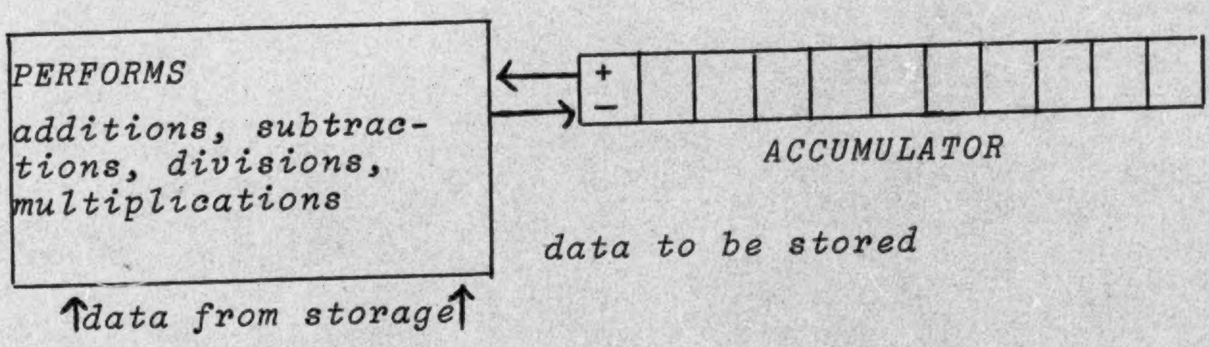

INPUT
(card reader)
CONTROL UNIT
ARITHMETIC UNIT
STORAGE UNIT
OUTPUT
(Printer)

Control Unit


Instruction
Operation
Address

The control unit is the part of the computer that determines which instruction is to be done next, decodes the operation to be done, and coordinates machine working with the programmed instruction. To begin, the instruction counter is set to show where the program begins.

Arithmetic Unit


PERFORMS
additions, subtractions, divisions, multiplications
+
–
ACCUMULATOR
data to be stored
data from storage

REFERENCES

1. School Mathematics Study Group, *Algorithms, Computation and Mathematics* [A.C. Vroman, Inc. (367 S. Pasadena Ave.), Pasadena, Calif. 91105, 1966].

2. Forsythe, A., Keenan, T., Organick, E., and Stenberg, W., *Computer Science: A First Course* (John Wiley & Sons, Inc., New York, 1969).

3. Forsythe, A., Keenan, T., Organick, E., and Stenberg, W., *Computer Science: A Primer* (John Wiley & Sons, Inc., New York, 1969).

DESCRIPTIVE TITLE	Synagraphic Computer-Mapping Program
CALLING NAME	SYMAP 5.12
INSTALLATION NAME	Computing Center, The Florida State University
AUTHOR(S) AND AFFILIATION(S)	Howard T. Fisher, Northwestern Technological Institute (overall design and mathematical model) Mrs. O.G. Benson, Northwestern University Computing Center (programming)
LANGUAGE	FORTRAN IV
COMPUTER	Originally developed on IBM 709; also on IBM 360/40 and IBM 7094 at Harvard University Available in EIN through The Florida State University CDC 6400
PROGRAM AVAILABILITY	Decks, listings, and documentation presently available from The Florida State University Computing Center
CONTACT	Ray Soller, Librarian, Computing Center, The Florida State University, Tallahassee, Fla. 32306 Tel.: (904) 599-3418

FUNCTIONAL ABSTRACT

The Synagraphic Computer-Mapping Program (SYMAP) produces maps that depict spatially disposed quantitative and qualitative information. Raw data of every kind (physical, social, economic, etc.) may be related, weighted, and aggregated in a graphic format by assigning values to the coordinate locations of data points or data zones. According to the application and desired representation of data, three basic types of mapping procedure may be specified: contour, conformant, or proximal.

CONTOUR—based on the use of *contour* lines, each of which represents a value remaining constant throughout its length. The map consists of closed curves that connect all points having the same numerical

continued

value. The value at each of the different levels of contour (where a single contour level will represent a particular data value) is determined by the program, according to the scale of the map and the range of the data. Between any two contour lines, a continuous variation is assumed. Therefore, the use of contour mapping should be restricted to the representation of continuously varying information, such as topography, rainfall, or population density.

CONFORMANT—based on the *conformance* to the boundaries of a data zone. This type of mapping is best suited to data for which the representation as a continuously varying surface is inappropriate owing to the significance of physical limits or boundaries. Each predefined data zone is assigned one data value and, depending on its numeric class (range), one representative character on the map itself. Local variation within the zone boundaries will not be apparent, but will, on the average, be correct.

PROXIMAL—based on *proximity* to a data point. In appearance, this type of map is similar to the conformant map. However, point information is used here to define the data zones. Each character location on the output map is assigned the value of the *nearest* data point, using nearest-neighbor techniques. Boundaries are then assumed along the lines where these values change. Then the mapping is carried out as in the conformant type.

REFERENCES

Robertson, J. C., "The SYMAP Programme for Computer Mapping," Cartographic 108–113 (Dec. 1967); taken from a report of the Select Committee on the Ordinance Survey of Scotland.

Fisher, M., "The Laboratory for Computer Graphics," Harvard Univ. Grad. School Design Suppl. (Summer 1967).

Shepard, D., "A Two-Dimensional Interpolation Function for Irregularly Spaced Data," Harvard Univ. Grad. School Design Lab. Computer Graphics (Feb. 1968); available from the Computing Center, The Florida State University. Deals with the subject of analyzing irregularly spaced data derived from a continuous surface. A method is developed for reconstructing the surface from the sampled data. This method is the main device used to generate the maps produced by SYMAP.

DESCRIPTIVE TITLE	Transportation Planning Package
CALLING NAME	TRAN/PLAN
INSTALLATION NAME	The Florida State University Computing Center
AUTHOR(S) AND AFFILIATION(S)	Charles E. Akers William T. Gaupin The Florida State University Computing Center
LANGUAGE	Primary Language: FORTRAN IV Secondary Language: COMPASS Assembly Language
COMPUTER	CDC 6400
PROGRAM AVAILABILITY	Programs generally available through the CDC users group. FSU will make them available to EIN members.
CONTACT	William T. Gaupin, Computing Center, The Florida State University, Tallahassee, Fla. 32306 Tel.: (904) 599-4770

FUNCTIONAL ABSTRACT

The Transportation Planning Package incorporates a set of analytical techniques that enable the user to estimate future transportation requirements and evaluate proposed systems. The collection of survey techniques, analysis method, and computer programs used by TRAN/PLAN have evolved over the past two decades through studies supported by State Highway Departments, U.S. Bureau of Public Roads, and the U.S. Department of Housing and Urban Development. The TRAN/PLAN programs perform the following functions, (1) process data accumulated from inventories or surveys of existing transportation facilities, (2) calibrate and analyze several parameters for a regional transportation network, and (3) predict future intraregional transportation demands. The accuracy of future estimates depends heavily upon the choice of a suitable model, effective intermediate analysis-of-program output, and iterative recalibration of model parameters. Reference 6 explains the overall purpose of transportation planning systems in greater detail.

continued

The TRAN/PLAN Package consists of the following subprograms[1].

(a)	RELINK	Relink nodes of network
(b)	NETWORK	Build and edit network
(c)	NETDVR	Network report generator
(d)	MINPATH	Build minimum paths
(e)	SKIM	Skim (summarize) tape of trees
(f)	COLLECT	Trip data collector
(g)	TRPBLDR	Build trip tables
(h)	ADDER	Add selected trip tables
(i)	SPLITER	Split trips by direction
(j)	UPDATE	Update trip volumes
(k)	MERGE	Merge trip tables
(l)	DEMERGE	Demerge trip tables
(m)	COMPRES	Compress or expand trip tables
(n)	SECTORS	Select sector of network
(o)	TRIPTAB	Trip-table summary generator
(p)	TRIPEND	Trip-end summary generator
(q)	GRAVITY	Gravity model
(r)	FRATAR	Fratar expansion
(s)	ASSIGN	Traffic assignment
(t)	CPACITY	Capacity restraint
(u)	MODSPLT	Modal split model
(v)	CAPRPT	Cross-reference historical-report generator
(w)	CPRSRPT	Historical-report generator
(x)	COMPARE	Statistical comparison of trip volumes
(y)	AVERAGE	Historical-record summary-report generator
(z)	TIMESUM	Time-table-report generator
(aa)	TRPFREQ	Trip frequency distribution

Restrictions which apply to all programs,

1. the maximum allowed zone (centroid) number is 650
2. the maximum allowed node number is 3000
3. the maximum number of tables (purposes) is 4
4. the maximum table number is 4
5. a stacked table may contain a maximum of 8 merged tables

REFERENCES

Akers, C.E. and Gaupin, W.T., "User Manual Transportation Programs for a CDC 6400 Computer," (The Florida State University Computing Center, November 1968; unpublished). Copies of this manual are available through the Librarian, The Florida State University Computing Center.

"Calibrating and Testing a Gravity Model for Any Size Urban Area," Bureau of Public Roads, Washington, D.C., October 1965.

continued

Manual Supplements for the Transportation Planning System, Highway and Local Government Department of South Austrailia, Adelaide, Austrailia, January 1966.

Martin, Brian V., "Minimum Path Algorithms for Transportation Planning," *Research Engineer*, December 1963.

"Modal Split," U.S. Department of Commerce, Bureau of Public Roads, Office of Planning, December 1966.

Shofer, R. and Goodyear, F., "Electronic Computer Applications in Urban Transportation Planning," Proceedings of the 22nd National Conference, Association for Computing Machinery Publication P-67, 1967.

Traffic Assignment Manual, Bureau of Public Roads, Washington, D.C., June 1964.

Transportation Planning System for the Control Data 3600 Computer, Data Centers Division of Control Data Corporation, Minneapolis, Minnesota, May 1965.

Fratar, T.J., "Vehicular Trip Distributions by Successvie Approximation," *Traffic Quarterly*, January 1954, pp. 53-65.

DESCRIPTIVE TITLE	Synagraphic Mapping
CALLING NAME	SYMAP
INSTALLATION NAME	The Pennsylvania State University Computation Center
AUTHOR(S) AND AFFILIATION(S)	Laboratory for Computer Graphics Harvard University Graduate School of Design Adapted by Larry Rich Jeffrey Simon Larry Sinkey Department of Architectural Engineering, The Pennsylvania State University
LANGUAGE	FORTRAN IV
COMPUTER	IBM System 360/67
PROGRAM AVAILABILITY	Decks, listings, and documentation presently available
CONTACT	Dr. Daniel Bernitt, EIN Technical Representative, 105 Computer Building, The Pennsylvania State University, University Park, Pa. 16802 Tel.: (814) 865-9527

FUNCTIONAL ABSTRACT

The Synagraphic Mapping program (SYMAP) produces maps that depict spatially disposed quantitative and qualitative information. Raw data of every kind (physical, social, economic, etc.) may be related, weighted, and aggregated in a graphic format by assigning values to the coordinate locations of data points or data zones. According to the application and desired representation of data, three basic types of mapping procedure may be specified: contour, conformant, or proximal.

CONTOUR
—based on the use of *contour* lines, each of which represents a value remaining constant throughout its length. The map consists of closed curves that connect all points having the same numerical value. The value at each of the different levels of contour

continued

(where a single contour level will represent a particular data value) is determined by the program, according to the scale of the map and the range of the data. Between any two contour lines, a continuous variation is assumed. Therefore, the use of contour mapping should be restricted to the representation of continuously varying information, such as topography, rainfall, or population density.

CONFORMANT
—based on the *conformance* to the boundaries of a data zone. This type of mapping is best suited to data for which the representation as a continuously varying surface is inappropriate owing to the significance of areal limits or boundaries. Each predefined data zone is assigned one data value and, depending on its numeric class (range), one representative character on the map itself. Local variation within the zone boundaries will not be apparent, but will, on the average, be correct.

PROXIMAL
—based on *proximity* to a data point. In appearance, this type of map is similar to the conformant map. However, point information is used here to define the data zones. Each character location on the output map is assigned the value of the *nearest* data point, using nearest-neighbor techniques. Boundaries are then assumed along the lines where these values change. Then, the mapping is carried out as in the conformant type.

While the contour type of map is mose often used and the easiest to produce, the conformant and proximal maps are often more helpful in the "soft" disciplines. Output is in the form of printed pages that, if the total map size exceeds the width of the computer printed page (13 in.), may easily be glued or pasted together to form a continuous map. Also included in the output is a histogram showing frequencies for given data levels, plus several optional features.

REFERENCES

Robertson, J.C., "The SYMAP Programme for Computer Mapping," Cartographic 108–113 (Dec. 1967); taken from a report of the Select Committee on the Ordinance Survey of Scotland.

Fisher, M., "The Laboratory for Computer Graphics," Harvard Univ. Grad. School Design Suppl. (summer 1967).

continued

Shepard, D., "A Two-Dimensional Interpolation Function for Irregularly Spaced Data," Harvard Univ. Grad. School Design Lab. Computer Graphics (Feb. 1968); available from the EIN Office for the cost of duplication and mailing. Deals with the subject of analyzing irregularly spaced data derived from a continuous surface. A method is developed for reconstructing the surface from the sampled data. This method is the main device used to generate the maps produced by SYMAP.

DESCRIPTIVE TITLE	Quick-Draw Graphics System
CALLING NAME	QDGS
INSTALLATION NAME	The Pennsylvania State University Computation Center
AUTHOR(S) AND AFFILIATION(S)	Jeff Raskin, The Pennsylvania State Univ. Computation Center; extended and revised by Graham Donaldson, The Pennsylvania State Univ. Computation Center
LANGUAGE	FORTRAN IV
COMPUTER	IBM System 360/67
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Daniel L. Bernitt, 105 Computer Building, The Pennsylvania State University, University Park, Pa. 16802 Tel.: (814) 865-9527

FUNCTIONAL ABSTRACT

QDGS is a set of FORTRAN subroutines that can be used to draw charts, graphs, diagrams, maps, or any other form of pictorial output. The graphic material is produced on the CalComp plotter and other graphic devices as they become available.

To use QDGS, a program is written in FORTRAN IV for the IBM System/360 computer. This program calls the QDGS subroutines; the subroutines produce output (in the form of cards or tape) that is later put into a special system to produce the actual graphical output. When the cards (or tape) are received from the dispatcher and the user has verified that the run was successful, he usually submits the cards to the dispatcher along with a plot request form.

The three concepts underlying the design of QDGS are (1) the description of an elementary picture in terms of coordinates, (2) the geometric transformation of pictures (into various positions and orientations), and (3) the building of complex pictures from elementary ones. There will be subroutines to do each of the three basic processes. In addition, there are special routines to draw axes for graphs and charts, to do lettering and numbering, and automatically to do the necessary task of punching the cards or

continued

writing the tape used in plotting (physically drawing on the plotting device). QDGS is intended as a nucleus about which the user's graphic routines, tailored to his own unique application, can be written.

DESCRIPTIVE TITLE	PL/I–FORMAC Interpreter
CALLING NAME	FORMAC
INSTALLATION NAME	The Pennsylvania State University Computation Center
AUTHOR(S) AND AFFILIATION(S)	R. Tobey J. Baker R. Crews P. Marks K. Victor International Business Machines Corporation
LANGUAGE	PL/I–FORMAC
COMPUTER	IBM 360/67
PROGRAM AVAILABILITY	Source code available on tape
CONTACT	Dr. Daniel Bernitt, EIN Technical Representative, Computation Center, 105 Computer Building, The Pennsylvania State University, University Park, Pa. 16802 Tel.: (814) 865-9527

FUNCTIONAL ABSTRACT

FORMAC is a system for carrying out formal manipulations on mathematical expressions. This allows for the use of analytic as well as numeric techniques. The most important capability of FORMAC is its accommodation of mathematical expressions as symbolic entities at execution time. For example, the execution of the FORMAC program segment

```
LET( A = X + Y ** 2;
     B = 2;
     C = A/B + 2.8);
```

may be interpreted as assigning the alphanumeric value $(X+Y^2)/2 + 2.8$ to the FORMAC variable C.

FORMAC enables the user to analyze expressions by identifying coefficients, common denominators, lead operators, and some characteristics of the operands. Constants can be factored, left in a rational form, or converted to real notation. New expressions can be synthesized by the simplification, expansion and substitution of

continued

terms, as well as finding the derivatives of functions. The user can specify functions completely or partially in addition to making use of the PL/I functions. Procedures are available for transfering arguments between PL/I and FORMAC program segments.

REFERENCES

Tobey, R., Baker, J., Crews, R., Marks, P., Victor, K., "PL/I-FORMAC Interpreter User's Reference Manual," IBM Publ. 360D 03.3.004 (1967). Copies of the User's Manual will be available through The Pennsylvania State University Computation Center to any user with a PSU account number.

DESCRIPTIVE TITLE	FORMAC Utility Program
CALLING NAME	FMACUT
INSTALLATION NAME	The Pennsylvania State University Computation Center
AUTHOR(S) AND AFFILIATION(S)	Robert Duquet Meteorology Department The Pennsylvania State University H.D. Knoble (assistance) The Pennsylvania State University Computation Center
LANGUAGE	PL/1—FORMAC
COMPUTER	IBM System 360/67
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Dr. Daniel L. Bernitt, EIN Technical Representative, Computation Center, 105 Computer Building, The Pennsylvania State University, University Park, Pa. 16802 Tel.: (814) 865-9527

FUNCTIONAL ABSTRACT

The FORMAC version available at The Pennsylvania State University is an interpretative system; i.e., FORMAC expressions are treated as character strings that are evaluated and executed at run time only. These character strings may be specified by reference to a PL/1 character-string variable. The PL/1 variable, in turn, may be constructed by execution of the PL/1 program in which the FORMAC statements are embedded. FMACUT merely capitalizes on the last-named feature. FORMAC statements are read by the preprocessed, precompiled, and prelink-edited program and are passed to the FORMAC package as character-string arguments. Note that this is made possible also because FORMAC variables need not (in fact *cannot*) be declared in the PL/1 program. Reference 1 should be consulted concerning more-detailed PL/1 FORMAC information.

FMACUT will execute FORMAC statements that are supplied as *input data*. It effectively divorces FORMAC from PL/1 insofar as the user is concerned. This has two advantages: (1) the user need have no knowledge of PL/1 to use FORMAC and (2) FORMAC is available without preprocessing, compilation, or link-editing (which

continued

saves approximately 50 seconds of run time for a null job on the IBM 360/67). The corresponding disadvantage is that PL/1 facilities and those (few) FORMAC statements involving a PL/1 direct interface are not available to the user.

REFERENCES

Tobey, R., Baker, J., Crews, R., Marks, P., and Victor, K., "PL/1 FORMAC Interpreter," IBM Publ. 360D 03.3.004 (1967).

DESCRIPTIVE TITLE	CalComp Plotter Subroutines
CALLING NAME	See specific subroutine
INSTALLATION NAME	The University of Iowa University Computer Center
AUTHOR(S) AND AFFILIATION(S)	CalComp Modified by the University of Iowa
LANGUAGE	FORTRAN IV
COMPUTER	IBM 360/65
PROGRAM AVAILABILITY	Available for use at the University of Iowa, but not for distribution
CONTACT	Mrs. Louise R. Levine, Program Librarian, Applications Programming, University Computer Center, The University of Iowa, Iowa City, Iowa 52240 Tel.: (319) 353-5580

FUNCTIONAL ABSTRACT

The CalComp plot routines are a series of subroutines used to obtain graphic output. They are designed for the following purposes:

1) providing values to scale data to fit the plotting area,
2) drawing identification symbols at plotted data points,
3) drawing connected lines between data points (when desired),
4) drawing and labeling axes at any desired orientation, and
5) selecting from a large number of characters, numbers, and special symbols, any of which can be drawn in any size and at any angular orientation anywhere on the plotting surface.

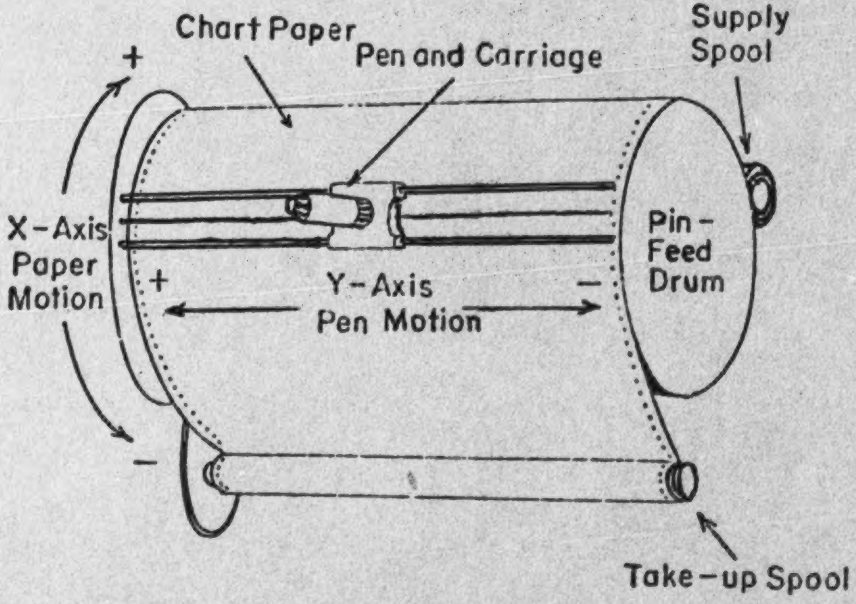
These routines have been written by CalComp and modified by the University of Iowa to run under OS on the IBM 360/65. Plotting is done off line. The user's plotter program will create a plot tape containing various commands used to drive a drum plotter (see fig. 1). Each of eight possible commands moves the pen 1/100 of an inch in one of the eight designated directions (see fig. 2). In addition there are commands to move

continued

the pen up and down. The plot tape is then taken to a smaller computer which drives the CalComp plotter (see fig. 3).

Access to the plot routines is by the FORTRAN "CALL" statement. Integer and floating point conventions must be followed (i.e. integer arguments must be used where the dummy argument specifies integer, likewise for floating point). All x and y coordinates used as arguments, must be expressed as floating point inches within actual page dimensions and in deflection from the (0.,0.) origin or the established reference point which is an established (0.,0.) origin. If the y coordinates are not within actual page dimensions in deflection from the reference point, that reference point will be destroyed and anything plotted after that will be erroneous. For orientation of the paper see fig. 1. All angles must be expressed in floating point degrees, with the positive sense counterclockwise. Character heights are specified as floating point inches. These heights should be greater than 0.07 but less than page size. Page size is 11 inches high and any reasonable length, since paper is continuous to a maximum length of 120 ft.

Figure 1



continued

Figure 2

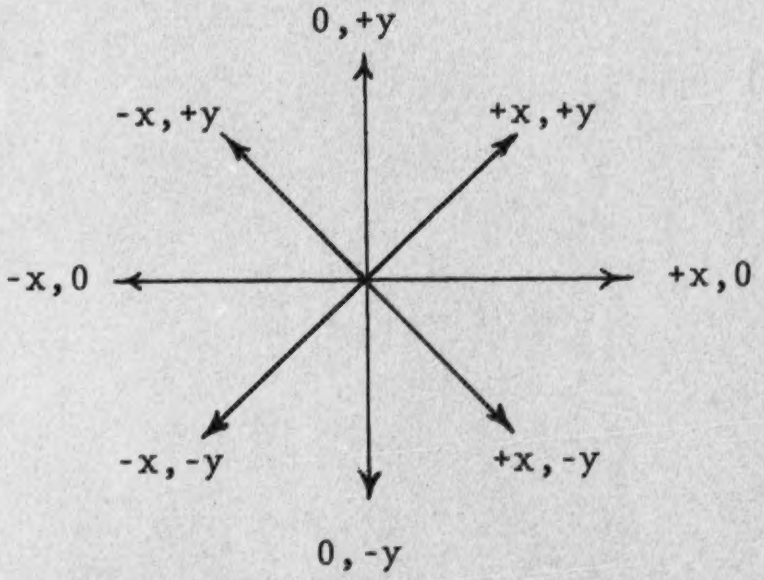
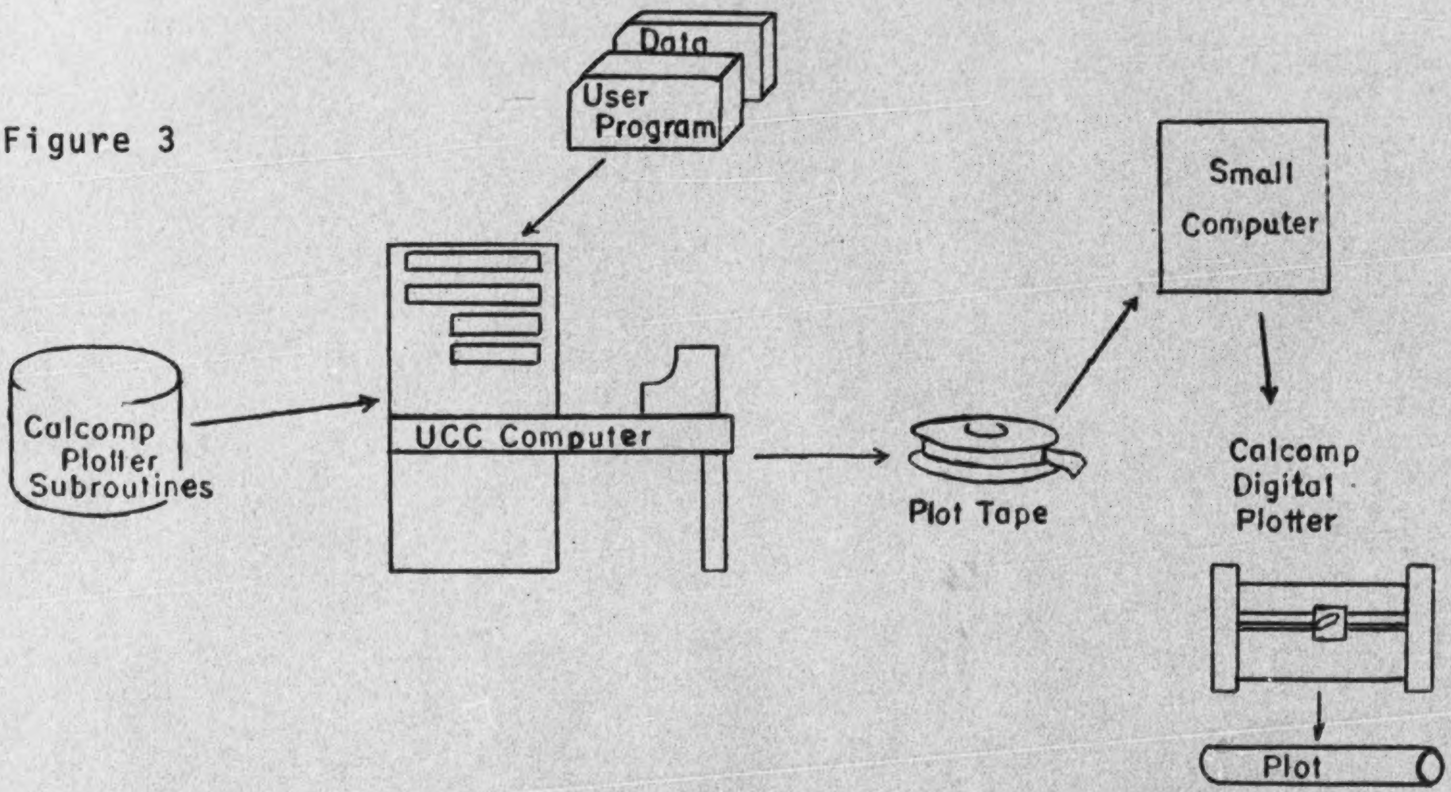

0,+y
-x,+y
+x,+y
-x,0
+x,0
-x,-y
+x,-y
0,-y

Figure 3


Data
User Program
Small Computer
Calcomp Plotter Subroutines
UCC Computer
Plot Tape
Calcomp Digital Plotter
Plot

REFERENCES

Plotter Subroutines for CalComp Digital Incremental Plotter (Iowa City, Ia.: Univ. Comp. Ctr., Univ. of Ia., 1970), Revised. Available from the EIN Office at the cost of reproduction and mailing.

DESCRIPTIVE TITLE	BEEF Data Processing Subroutines
CALLING NAME	BEEFDP
INSTALLATION NAME	The University of Notre Dame Computing Center
AUTHOR(S) AND AFFILIATION(S)	Westinghouse Electric Corporation UNIVAC Division of Sperry Rand Corporation
LANGUAGE	SLEUTH II/FORTRAN IV
COMPUTER	UNIVAC 1107
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Elizabeth Hutcheson, EIN Technical Representative, Computing Center, University of Notre Dame, Notre Dame, Ind. 46556 Tel.: (219) 283-7784

FUNCTIONAL ABSTRACT

The BEEF Data Processing library is a set of subroutines supplied by UNIVAC to "enhance FORTRAN's abilities as a commercial processor." This enrichment is in the form of subroutines for whole word data movement, character and field movement, supplementary formatting, decision making with FORTRAN arrays, data conversion, report generation and control, I/O control, sorting, and compatability with EAM/EDP (Electronic Accounting Machine/Electronic Data Processing) equipment.

REFERENCE

UNIVAC 1107 BEEF Data Processing Manual, (UP-3985), (UNIVAC Division of Sperry Rand Corp., N.Y.). Copies of this manual are available through the local UNIVAC representative or through the EIN Office at the cost of reproduction and mailing.

DESCRIPTIVE TITLE	UNIVAC 1107 Linear Programming Package
CALLING NAME	LP1107
INSTALLATION NAME	The University of Notre Dame Computing Center
AUTHOR(S) AND AFFILIATION(S)	UNIVAC Division of Sperry Rand
LANGUAGE	Sleuth II/FORTRAN IV
COMPUTER	UNIVAC 1107
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Elizabeth Hutcheson, EIN Technical Representative, Computing Center, Univ. of Notre Dame, Notre Dame, Ind. 46556 Tel.: (219) 283-7784

FUNCTIONAL ABSTRACT

LP1107 is a generalized program for the solution of linear programming problems. More specifically, the LP1107 command structure is a mathematical programming control language; used to prepare a control sequence for a specific programming job.

Either the dual or simplex algorithm, or both, can be used. LP1107 incorporates a true programming language with logical capabilities and full macro capabilities. If a group of commands is being used repetitively, the user may incorporate it into his macro command library and issue the macro command instead. Single and double precision arithmetic are included.

REFERENCES

UNIVAC Systems Programming Library Services, *UP3897 1107LP User's Manual (preliminary)*, (UNIVAC, Sperry Rand Building, New York, 1964).

Aronofsky, J.S., "Growing Applications of Linear Programming with a View Towards Understanding Principles on Which Future LP Systems Must be Based," *Communications of the ACM*, (June 1964).

Copies of these references are available through the local UNIVAC representatives or from the EIN Office at the cost of reproduction and mailing.

DESCRIPTIVE TITLE	Plot
CALLING NAME	PLOT
INSTALLATION NAME	Office of Data Analysis Research Educational Testing Service (ETS)
AUTHOR(S) AND AFFILIATION(S)	University of Michigan
LANGUAGE	MAP
COMPUTER	IBM 360/65
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Mr. Ernest Anastasio, Off. of Data Anal. Research, Educational Testing Service Rosedale Road, Princeton, N.J. 08540 Tel.: (609) 921-9000 ext. 2552

FUNCTIONAL ABSTRACT

A flexible plot routine originally written for the 7090 at the University of Michigan and adapted to their system, has been modified for the operating system at ETS. Although the program is written in assembly language, the necessary linkage to FORTRAN and FORTRAN I/O has been written. It is intended to be called, therefore, by a FORTRAN program and will produce plots in the normal output stream as determined by the various calls, to different entry points in the subroutine. The standard approach is to call PLOT1, PLOT2, PLOT3, and PLOT4 in that order. PLOT1 sets up the information required to construct the graph, PLOT2 prepares the grid and sets up the information required by PLOT3 to place points in the graph, PLOT3 places the plotting characters at the specified points in the graph, and PLOT4 prints the completed graph with values along the X and Y axes and a centered vertical label down the left side of the graph.

DESCRIPTIVE TITLE	Simulation Package for University Research and Training
CALLING NAME	SPURT
INSTALLATION NAME	Vogelback Computing Center Northwestern University
AUTHOR(S) AND AFFILIATION(S)	Gustave J. Rath Department of Industrial Engineering and Management Sciences Martin Goldberg Leonard Weiner Northwestern University
LANGUAGE	CDC FORTRAN IV
COMPUTER	CDC 6400
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Lorraine Borman, EIN Technical Representative, Vogelback Computing Center, Northwestern University, 2129 Sheridan Road, Evanston, Ill. 60201 Tel.: (312) 492-3682

FUNCTIONAL ABSTRACT

SPURT is a comprehensive package of USASI Standard FORTRAN routines that are designed for use in simulation modelling. These useful routines, ranging from simple to complex, enable the average FORTRAN programmer to employ simulation techniques without having to learn the semantic and syntactic rules of a new programming language.

The SPURT package is made up of six main parts.

I. *CLOCK Generation—SPURT1*
II. *Stochastic Generators—SPURT2*
III. *Statistical Computations—SPURT3*
IV. *Analog Simulators—SPURT4*
V. *List-Processing and Queue-Manipulation—SPURT5*
VI. *Matrix and Graphical Output—SPURT6*

continued

The following is a listing and brief discussion of the various subroutines contained in each of the six SPURT parts.

CLOCK Generation—SPURT1: to implement discrete-time simulation models; to cause events to occur in the proper time sequence

The CLOCK subroutine consists basically of two lists:

Master Time List—contains events scheduled to happen in the future

Master Time Queue—contains events that could not take place at the time when they were scheduled to and, therefore, have been rescheduled; i.e., they have been blocked and are waiting in a queue.

Events can be stored on either list.

CLOCK recognizes two basic kinds of events:

Exogenous—those that are *external* to the user's routine; these are read from Data Cards by the CLOCK

Endogenous—those that are *internal* to the user's routine; these are generated dynamically and then are maintained by the CLOCK

Stochastic Generators—SPURT2: to generate samples from various probability distributions and to calculate sample values

Subroutine	*Usage*
STOGN1	Permits sampling from a discrete empirical probability distribution defined by the user
STOGN2	Enables the user to approximate a continuous distribution by means of a piecewise linear distribution
UNIFRM	Permits the user to sample real values from a uniform distribution in a defined interval
RANDIN	Provides a uniform distribution of integers in a defined interval
NORMAL	Allows the user to obtain a random sample from a normal distribution with given mean and standard deviation
NEGEXP	Permits the user to obtain a random sample from the negative exponential distribution
POISSN	Provides the user with a random sample from the Poisson distribution

continued

ERLANG	Provides a random sample from the Erlang distribution
DISCRT	Permits sampling from a step function describing a discrete cumulative distribution of integer values
LINEAR	Provides the user with a random sample from a cumulative distribution that is obtained by linear interpolation in a nonequidistant table of real values
DRAW	Provides a boolean value of TRUE or FALSE
RANPER	Generates a uniformly distributed, random permutation of the integers 1, 2, ···, M

Statistical Computations—SPURT3: to calculate statistical parameters and histograms of data arrays

Subroutine	*Usage*
STIX1 STIX2 STIX3	Three interrelated subroutines to accumulate and print out a frequency table and to produce a CalComp plot of a normalized histogram of the table
STIX4	Evaluates the mean, standard deviation, maximum value, and minimum value of an array of real numbers
STIX5	Evaluates the correlation coefficient between two arrays of real numbers
STIX6	Ranks an array of real numbers and produces the median and range of the data within the array
STIX7	Produces a statistical description of the data found in an array, including the sample size, mean, standard deviation, standard error, minimum and maximum values, range, and a printed histogram plot

Analog Simulators—SPURT4: to enable the simulation of analog-computer problems on a digital computer

Subroutine	*Usage*
ANALOG SECND	These two subroutines make it possible to obtain output similar to a hybrid computer

continued

List-Processing and Queue-Manipulation—SPURT5: lists are mXn arrays; entries in lists are mX1 arrays.

Subroutine	*Usage*
ADFIFO	Adds an entry at the bottom of the list; it can be removed only after all the elements presently on the list are gone (builds first-in last-out list)
ADLIFO	Adds an entry at the top of the list; it will be removed before any other entry presently on the list (builds last-in first-out list)
REMOVE	Removes the top (or first) entry from a list
PURGE	Destroys the contents of a list
DISPL	Prints the contents of a list

Additional subroutines in SPURT5 provide the capability to rank lists and to delete or to insert entries into lists.

Matrix and Graphical Output—SPURT6: output is facilitated through printing and graphical output

Subroutine	*Usage*
OUT	Prints out a square matrix with column and row headings
NSOUT	Prints out a nonsquare matrix with column and row headings
GRAPH	Produces two-dimensional graphs of plots, using a CalComp plotter

REFERENCES

1. Rath, G.J., "Description of the Simulation Package for University Research and Teaching," Northwestern Univ. Vogelback Computing Ctr., SPURT, Rev. E (1968; unpublished).

DESCRIPTIVE TITLE	TEXT360: A System for Producing Manuals
CALLING NAME	TEXT360
INSTALLATION NAME	Washington University Computing Facilities
AUTHOR(S) AND AFFILIATION(S)	S.L. Reed, International Business Machines
LANGUAGE	PL/I and Assembly Language
COMPUTER	IBM/360 under OS
PROGRAM AVAILABILITY	Proprietary; available for use but not for distribution
CONTACT	J. Philip Miller, Computing Facilities, Box 1152, Washington University, St. Louis, Mo. 63130 Tel.: (314) 863-0100 ext. 3141

FUNCTIONAL ABSTRACT

TEXT360 is a text-processing system with data-entry, data-updating, and page-formatting capabilities. The system, which runs under Operating System/360, consists of a main processor and several peripheral programs. Input to the system is free-form and is produced on the IBM 029 Card Punch. Output is camera-ready and is produced on the IBM 1403 Printer. The formatting capabilities of TEXT360 permit the insertion, deletion, and replacement of characters, words, lines, and groups of lines. In addition, blocks of text can be moved from one part of a document to another. One-column and two-column page format can be produced. Routine functions include hyphenation, line justification, column heading, and indentations. More complex functions include the generation of horizontal and vertical ruling for tables and figures. The program also allows the user to specify that related material (e.g., a table) is to be kept together, i.e., not split between columns or pages.

The four phases of the TEXT360 Formatting Processor, the TEXT360 Spelling Dictionary Update Program, the TEXT90-to-TEXT360 Master File Conversion Program and the Print/Punch Utility Program are

continued

written in the PL/1 Language, supplemented by four small assembler-language routines used for character-set mapping. The TEXT360 Prescan and Peripheral Print Programs are written in System/360 assembler language.

TEXT360 is essentially the program used by IBM for the production of the SRL manuals.

REFERENCES

Reed, S.L., *TEXT360*, Contributed Program Library Documentation, Program 360D-29.4.001 (White Plains, N. Y.: IBM, 1967).

000 0080

000 0080

DESCRIPTIVE TITLE	Synagraphic Mapping Program
CALLING NAME	SYMAP
INSTALLATION NAME	Washington University Computing Facilities
AUTHOR(S) AND AFFILIATION(S)	Harvard University Graduate School of Design
LANGUAGE	FORTRAN IV
COMPUTER	IBM 360/50
PROGRAM AVAILABILITY	Proprietary; available for use but not distribution
CONTACT	Dr. Charles Drebes, Manager, Scientific Data Processing, Computing Facilities, Box 1098, Washington University, St. Louis, Mo. 63130 Tel.: (314) 863-0100 ext. 3141

FUNCTIONAL ABSTRACT

SYMAP is a computer program designed to allow city planners, geographers and others to produce low cost graphic displays of spatial patterns using standard computer line printers, by producing maps which graphically depict spatially disposed quantitative and qualitative information. It is suited to a broad range of applications, and is provided with numerous options to meet widely varying requirements. Raw data of every kind (physical, social, economic, etc.) when given to the computer may be related, manipulated, weighted, and aggregated in any manner desired. By assigning values to the coordinate locations of data points or data zones, one or more of three types of map may be produced, as specified by the user: conformant (choropleth), contour, and proximal. Potential applications are independent of the scale at which one wishes to display data. Studies (at other universities) have included a living cell, land parcels, blocks, tracts, towns, states, and continents. In each case, a common factor was the spatial distribution of a variable and a need to display the patterns associated with this distribution.

REFERENCES

Reference Manual for Synagraphic Computer Mapping—"SYMAP" (Cambridge, Mass.: Harvard Univ. Grad. Sch. of Design, Comp. Graph. Lab.).

000 0080

DESCRIPTIVE TITLE	Bucharest Sort a List into Ascending Order
CALLING NAME	ABSRT
INSTALLATION NAME	Iowa State University Computation Center
AUTHOR(S) AND AFFILIATION(S)	Iowa State University Computation Center
LANGUAGE	FORTRAN IV
COMPUTER	IBM System 360/65
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Dr. Clair G. Maple, Director, Computation Ctr., Iowa State University, Ames, Ia. 50010 Tel.: (515) 294-3402

FUNCTIONAL ABSTRACT

ABSRT is a subroutine to sort the contents of an array into ascending order [defined as $A(N) \leq A(N+1)$]. It is written for a list of real numbers; however, with a slight revision in a declaration statement within the program deck, it may be easily altered to sort a list of integers or double-precision real numbers.

A formal description of the two-way, internal, single-phase merge method may be found in the literature.[1] Briefly described, the procedure considers the input list to be composed of a number of shorter sublists that are already ordered. (Initially, every element in the list comprises a sublist.) Half of the sublists are stored in standard order, forward from A(1), and the other half are stored backwards from A(N)—i.e., in A(N), A(N–1), A(N–2), etc. The first pass through the input list merges pairs of these sublists, taking one from each end of the input list, and stores each newly merged pair in the output list. The output list is stored in A(N+1) through A(2N), and, like the input list, is made up of sublists stored forward from A(N+1) and backward from A(2N). On successive passes through the list, the input and output areas are switched, and consecutively, the number of sublists is halved. When only one sublist remains, the list is in order. At that time, a check is made to determine if the sorted list is in the first or second half of the array; if necessary, a final pass is made to put the list back into the first half, A(1) through A(N).

REFERENCES

1. Iverson, K.E., *A Programming Language* (John Wiley & Sons, Inc., New York, 1962), pp.206–211.

DESCRIPTIVE TITLE	Bucharest Sort a List into Descending Order
CALLING NAME	DBSRT
INSTALLATION NAME	Iowa State University Computation Center
AUTHOR(S) AND AFFILIATION(S)	Iowa State University Computation Center
LANGUAGE	FORTRAN IV
COMPUTER	IBM System 360/65
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Dr. Clair G. Maple, Director, Computation Ctr., Iowa State University, Ames, Ia. 50010 Tel.: (515) 294-3402

FUNCTIONAL ABSTRACT

DBSRT is a subroutine to sort the contents of an array into descending order [defined as $A(N) \geq A(N+1)$]. It is written for a list of real numbers; however, with a slight revision in a declaration statement within the program deck, it may be easily altered to sort a list of integers or double-precision real numbers.

The method used is identical to that described in the entry for subroutine ABSRT (EIN No. 000 0084), except, of course, for the reversed ordering of the elements within the mentioned sublists. For convenience, the reference from which a formal description of the sorting algorithm (the two-way, internal, single-phase merge method) may be obtained is listed below.

REFERENCES

Iverson, K.E., *A Programming Language* (John Wiley & Sons, Inc., New York, 1962), pp. 206–211.

DESCRIPTIVE TITLE	Reformat into Either BCD or EBCDIC
CALLING NAMES	EDP C 005 E (from BCD to EBCDIC) EDP C 005 B (from EBCDIC to BCD)
INSTALLATION NAME	Computation Center, The City College of The City University of New York
AUTHOR(S) AND AFFILIATION(S)	A. J. Cohen G. W. Elder Computation Center, The City College of The City University of New York
LANGUAGE	COBOL
COMPUTER	IBM 7040, IBM 360/50 O.S., IBM 360/30 D.O.S.
PROGRAM AVAILABILITY	Deck and listing presently available
CONTACT	R. M. Lobou, Operations Supervisor, Computation Center, The City College of The City University of New York, New York, N.Y. 10031 Tel.: (212) 621-2374

FUNCTIONAL ABSTRACT

This program is designed to reformat data and to convert decks from BCD to EBCDIC or from EBCDIC to BCD. It can be used on any second- or third-generation equipment that supports COBOL. In addition to conversion, there exists an option to reorganize the data completely. Interspersed gangpunching is also possible. That is, constants —such as dates or program numbers—may be emitted into the output data.

The input data may appear in mixed-character configuration and yet yield decks in the desired character set. That is, each input card can be punched in either BCD or EBCDIC and will be punched as output in the specified punch configuration.

000 0091

DESCRIPTIVE TITLE	Matrix Operations
CALLING NAME	AES 106
INSTALLATION NAME	The Michigan State University Computer Laboratory, Applications Programming
AUTHOR(S) AND AFFILIATION(S)	William L. Ruble Richard J. Martz Agricultural Experimental Station Michigan State University
LANGUAGE	3600 COMPASS (closed subroutines callable from FORTRAN 600 or COMPASS)
COMPUTER	CDC 3600
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Mr. Anders Johanson, 324 Computer Center, Michigan State University, East Lansing, Mich. 48823 Tel.: (517) 355-4684

FUNCTIONAL ABSTRACT

This is a package of several subroutines for double-precision matrix manipulation.

REFERENCES

Michigan State University Agricultural Experimental Station, *AES MISC Series Description Numbers 106C, 106D, 106E, 106F, 106I, 106J, 106K, 106L*. Mimeographed descriptions available through Mr. Johanson, above, or from EIN at the cost of reproduction.

DESCRIPTIVE TITLE	Page Plotter Using Line Printer
CALLING NAME	LPLOT
INSTALLATION NAME	The University of Iowa University Computer Center
AUTHOR(S) AND AFFILIATION(S)	James J. Hurt, for 7044 Version Wayne Robinson, conversion for 360 University Computer Center The University of Iowa
LANGUAGE	FORTRAN IV and Assembler Language for IBM 360/65
COMPUTER	IBM 360/65
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Mrs. Louise R. Levine, Program Librarian, University Computer Center, The Univ. of Iowa, Iowa City, Iowa 52240 Tel.: (319) 353-5580

FUNCTIONAL ABSTRACT

The subroutine LPLOT produces on the line printer a plot of from one to eight ordinate arrays versus one abscissa array on the same plot. A plot summary of the maximums, minimums, and scale per inch for the abscissa and ordinates is printed below the plot.

DESCRIPTIVE TITLE	UCSB On-Line System
CALLING NAME	OLS
INSTALLATION NAME	University of California Santa Barbara UCSB Computer Center
AUTHOR(S) AND AFFILIATION(S)	Dr. David Harris, Principal Investigator Ron Stoughton, Principal Programmer UCSB Computer Research Lab.
LANGUAGE	Basic Assembler Language
COMPUTER	IBM System 360/75
PROGRAM AVAILABILITY	Listing of programs may be obtained by special arrangement with the Computer Research Lab.
CONTACT	Glenn Davis, OLS Manager, Computer Center, Univ. of Cal., Santa Barbara, Cal. 93106 Tel.: (805) 961-2462

FUNCTIONAL ABSTRACT

The UCSB On-Line System (OLS) provides the capability for sophisticated mathematical analysis for use in solving problems where human interaction is either necessary or desired.

OLS accepts both real and complex numbers (scalars) as operands as well as lists of such numbers (vectors). Operations performed on scalars produce scalar results, which can be numerically displayed; operations on vectors produce vector results (the specified operation being performed on each component), and results of computation can be displayed either numerically or graphically. Operands can be stored and used as required. Operators include sine, cosine, logarithm, and exponentiation; and each is executed with a single button push. Facility is provided for interaction between operands of different types (e.g. vectors and scalars). In addition, a limited set of operations manipulate integers used in subscripting.

Additional features are provided to support OLS's basic mathematical capability. Although OLS normally executes each

continued

button as it is pushed, a button sequence may be defined, named, and saved for later execution. Convenient means are provided for editing such sequences. Lists of buttons to be executed can include programmed pauses, allowing manual and programmed activity to be interfaced; as well as branching based upon results of computation. Messages can be composed of alphameric, Greek, and special characters, and displayed. Those characters not specifically provided by OLS may be designed by the user and stored, and then are available for use. A collection of button lists and user-created characters is referred to as a "system". Systems are named and can be permanently stored and later retrieved. Portions of systems may be transferred between systems, and systems may be transferred between users. Sets of scalars and vectors may also be named, permanently stored, and later retrieved.

Apart from OLS's mathematical capability, a recent development provides the ability to create and edit a "deck" of cards and submit it for execution in an OS partition. Operations on string, record and file levels are provided. Data-sets residing on any disk pack within the installation may be fetched, examined, modified, and submitted for execution. Work continues in the general field and further developments are expected.

DESCRIPTIVE TITLE	General Purpose Simulation System
CALLING NAME	GPSS
INSTALLATION NAME	Washington University Computing Facilities
AUTHOR(S) AND AFFILIATION(S)	IBM Application Program
LANGUAGE	360 Assembly Language
COMPUTER	IBM 360/50
PROGRAM AVAILABILITY	Proprietary; available for use at Washington University. Available for distribution from IBM.
CONTACT	Dr. C.B. Drebes, Mgr., Scientific Data Processing, Computing Facilities, Box 1098, Washington University, St. Louis, Mo. 63130 Tel.: (314) 863-0100 ext. 3141

FUNCTIONAL ABSTRACT

GPSS is a transaction-oriented language designed for conducting evaluations and experiments concerning the behavior of systems, methods and processes. It has a modular structure which permits "transactions" to flow through the system, where their interactions can be observed and modified. A "clock" is maintained by which events are either scheduled to occur or else determined by one of the eight random number generators provided. Information can be obtained regarding sequencing of operations, scheduling and allocation rules, inventories, queuing disciplines, machine failures, etc. In general, various trade-offs between cost and performance can be studied.

REFERENCES

General Purpose Simulation System/360, Introductory User's Manual (IBM Manual GH20-0304-0), (White Plains, N.Y.: IBM Corp. Tech. Pub. Depart., 1969).

General Purpose Simulation System/360, User's Manual (IBM Manual GH20-0326-3), (White Plains, N.Y.: IBM Corp. Tech. Pub. Depart., 1969).

General Purpose Simulation System/360, OS (360A-CS-17X) Operator's Manual (IBM Manual H20-0311), (White Plains, N.Y.: IBM Corp. Tech. Pub. Depart., 1969).

DESCRIPTIVE TITLE	A Primal-Dual Transportation Algorithm
CALLING NAME	TRANSPRT
INSTALLATION NAME	Washington University Computing Facilities
AUTHOR(S) AND AFFILIATION(S)	D.E. Burlingame Washington University Computing Facilities
LANGUAGE	FORTRAN IV
COMPUTER	IBM 360/50
PROGRAM AVAILABILITY	Deck and listing presently available
CONTACT	Dr. C.B. Drebes, Mgr., Scientific Data Processing, Computing Facilities, Box 1098, Washington University, St. Louis, Mo. 63130 Tel.: (314) 863-0100 ext. 3141

FUNCTIONAL ABSTRACT

TRANSPRT solves a standard transportation problem using the primal-dual transportation algorithm. This program solves a model in which the objective is to "transport" a single commodity from various origins to different destinations at a minimum total shipping cost. The availability at each origin, the demand at each destination, and the cost to ship one unit of the project from any origin to any destination are required inputs to this model. This model can be applied to certain other types of industrial or business problems that have nothing to do with shipping. Personnel assignment, machine assignment, product and inventory scheduling are a few such applications.

The program utilizes integer arithmetic and requires integer input. If the sum of the origin availabilities is not equal to the sum of the destination requirements, an artificial origin or destination is set up, with zero costs, to handle the excess. The program will handle up to 50 origins and 150 destinations, using 94K bytes of core storage.

REFERENCES

Hadley, G., *Linear Programming* (Reading, Mass: Addison-Wesley Pub. Co., Inc., 1962).

Earth Sciences

EDUCOM

EDUCATIONAL INFORMATION NETWORK

000 0039

DESCRIPTIVE TITLE	Synagraphic Computer-Mapping Program
CALLING NAME	SYMAP 5.12
INSTALLATION NAME	Computing Center, The Florida State University
AUTHOR(S) AND AFFILIATION(S)	Howard T. Fisher, Northwestern Technological Institute (overall design and mathematical model) Mrs. O.G. Benson, Northwestern University Computing Center (programming)
LANGUAGE	FORTRAN IV
COMPUTER	Originally developed on IBM 709; also on IBM 360/40 and IBM 7094 at Harvard University Available in EIN through The Florida State University CDC 6400
PROGRAM AVAILABILITY	Decks, listings, and documentation presently available from The Florida State University Computing Center
CONTACT	Ray Soller, Librarian, Computing Center, The Florida State University, Tallahassee, Fla. 32306 Tel.: (904) 599-3418

FUNCTIONAL ABSTRACT

The Synagraphic Computer-Mapping Program (SYMAP) produces maps that depict spatially disposed quantitative and qualitative information. Raw data of every kind (physical, social, economic, etc.) may be related, weighted, and aggregated in a graphic format by assigning values to the coordinate locations of data points or data zones. According to the application and desired representation of data, three basic types of mapping procedure may be specified: contour, conformant, or proximal.

CONTOUR—based on the use of *contour* lines, each of which represents a value remaining constant throughout its length. The map consists of closed curves that connect all points having the same numerical

continued

value. The value at each of the different levels of contour (where a single contour level will represent a particular data value) is determined by the program, according to the scale of the map and the range of the data. Between any two contour lines, a continuous variation is assumed. Therefore, the use of contour mapping should be restricted to the representation of continuously varying information, such as topography, rainfall, or population density.

CONFORMANT—based on the *conformance* to the boundaries of a data zone. This type of mapping is best suited to data for which the representation as a continuously varying surface is inappropriate owing to the significance of physical limits or boundaries. Each predefined data zone is assigned one data value and, depending on its numeric class (range), one representative character on the map itself. Local variation within the zone boundaries will not be apparent, but will, on the average, be correct.

PROXIMAL—based on *proximity* to a data point. In appearance, this type of map is similar to the conformant map. However, point information is used here to define the data zones. Each character location on the output map is assigned the value of the *nearest* data point, using nearest-neighbor techniques. Boundaries are then assumed along the lines where these values change. Then the mapping is carried out as in the conformant type.

REFERENCES

Robertson, J. C., "The SYMAP Programme for Computer Mapping," Cartographic 108–113 (Dec. 1967); taken from a report of the Select Committee on the Ordinance Survey of Scotland.

Fisher, M., "The Laboratory for Computer Graphics," Harvard Univ. Grad. School Design Suppl. (Summer 1967).

Shepard, D., "A Two-Dimensional Interpolation Function for Irregularly Spaced Data," Harvard Univ. Grad. School Design Lab. Computer Graphics (Feb. 1968); available from the Computing Center, The Florida State University. Deals with the subject of analyzing irregularly spaced data derived from a continuous surface. A method is developed for reconstructing the surface from the sampled data. This method is the main device used to generate the maps produced by SYMAP.

DESCRIPTIVE TITLE	Synagraphic Mapping
CALLING NAME	SYMAP
INSTALLATION NAME	The Pennsylvania State University Computation Center
AUTHOR(S) AND AFFILIATION(S)	Laboratory for Computer Graphics Harvard University Graduate School of Design Adapted by Larry Rich Jeffrey Simon Larry Sinkey Department of Architectural Engineering, The Pennsylvania State University
LANGUAGE	FORTRAN IV
COMPUTER	IBM System 360/67
PROGRAM AVAILABILITY	Decks, listings, and documentation presently available
CONTACT	Dr. Daniel Bernitt, EIN Technical Representative, 105 Computer Building, The Pennsylvania State University, University Park, Pa. 16802 Tel.: (814) 865-9527

FUNCTIONAL ABSTRACT

The Synagraphic Mapping program (SYMAP) produces maps that depict spatially disposed quantitative and qualitative information. Raw data of every kind (physical, social, economic, etc.) may be related, weighted, and aggregated in a graphic format by assigning values to the coordinate locations of data points or data zones. According to the application and desired representation of data, three basic types of mapping procedure may be specified: contour, conformant, or proximal.

CONTOUR
—based on the use of *contour* lines, each of which represents a value remaining constant throughout its length. The map consists of closed curves that connect all points having the same numerical value. The value at each of the different levels of contour

continued

(where a single contour level will represent a particular data value) is determined by the program, according to the scale of the map and the range of the data. Between any two contour lines, a continuous variation is assumed. Therefore, the use of contour mapping should be restricted to the representation of continuously varying information, such as topography, rainfall, or population density.

CONFORMANT
—based on the *conformance* to the boundaries of a data zone. This type of mapping is best suited to data for which the representation as a continuously varying surface is inappropriate owing to the significance of areal limits or boundaries. Each predefined data zone is assigned one data value and, depending on its numeric class (range), one representative character on the map itself. Local variation within the zone boundaries will not be apparent, but will, on the average, be correct.

PROXIMAL
—based on *proximity* to a data point. In appearance, this type of map is similar to the conformant map. However, point information is used here to define the data zones. Each character location on the output map is assigned the value of the *nearest* data point, using nearest-neighbor techniques. Boundaries are then assumed along the lines where these values change. Then, the mapping is carried out as in the conformant type.

While the contour type of map is mose often used and the easiest to produce, the conformant and proximal maps are often more helpful in the "soft" disciplines. Output is in the form of printed pages that, if the total map size exceeds the width of the computer printed page (13 in.), may easily be glued or pasted together to form a continuous map. Also included in the output is a histogram showing frequencies for given data levels, plus several optional features.

REFERENCES

Robertson, J.C., "The SYMAP Programme for Computer Mapping," Cartographic 108–113 (Dec. 1967); taken from a report of the Select Committee on the Ordinance Survey of Scotland.

Fisher, M., "The Laboratory for Computer Graphics," Harvard Univ. Grad. School Design Suppl. (summer 1967).

continued

Shepard, D., "A Two-Dimensional Interpolation Function for Irregularly Spaced Data," Harvard Univ. Grad. School Design Lab. Computer Graphics (Feb. 1968); available from the EIN Office for the cost of duplication and mailing. Deals with the subject of analyzing irregularly spaced data derived from a continuous surface. A method is developed for reconstructing the surface from the sampled data. This method is the main device used to generate the maps produced by SYMAP.

DESCRIPTIVE TITLE	Synagraphic Mapping Program
CALLING NAME	SYMAP
INSTALLATION NAME	Washington University Computing Facilities
AUTHOR(S) AND AFFILIATION(S)	Harvard University Graduate School of Design
LANGUAGE	FORTRAN IV
COMPUTER	IBM 360/50
PROGRAM AVAILABILITY	Proprietary; available for use but not distribution
CONTACT	Dr. Charles Drebes, Manager, Scientific Data Processing, Computing Facilities, Box 1098, Washington University, St. Louis, Mo. 63130 Tel.: (314) 863-0100 ext. 3141

FUNCTIONAL ABSTRACT

SYMAP is a computer program designed to allow city planners, geographers and others to produce low cost graphic displays of spatial patterns using standard computer line printers, by producing maps which graphically depict spatially disposed quantitative and qualitative information. It is suited to a broad range of applications, and is provided with numerous options to meet widely varying requirements. Raw data of every kind (physical, social, economic, etc.) when given to the computer may be related, manipulated, weighted, and aggregated in any manner desired. By assigning values to the coordinate locations of data points or data zones, one or more of three types of map may be produced, as specified by the user: conformant (choropleth), contour, and proximal. Potential applications are independent of the scale at which one wishes to display data. Studies (at other universities) have included a living cell, land parcels, blocks, tracts, towns, states, and continents. In each case, a common factor was the spatial distribution of a variable and a need to display the patterns associated with this distribution.

REFERENCES

Reference Manual for Synagraphic Computer Mapping—"SYMAP" (Cambridge, Mass.: Harvard Univ. Grad. Sch. of Design, Comp. Graph. Lab.).

DESCRIPTIVE TITLE	Mineral Identification
CALLING NAME	MINERAL
INSTALLATION NAME	Kiewit Computation Center Dartmouth College
AUTHOR(S) AND AFFILIATION(S)	Michael Hebb Kiewit Computation Center Dartmouth College
LANGUAGE	Dartmouth BASIC
COMPUTER	GE-635
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	A. Kent Morton, EIN Technical Representative, Kiewit Computation Center, Dartmouth College, Hanover, N.H. 03755 Tel.: (603) 646-2864

FUNCTIONAL ABSTRACT

MINERAL was written as an aid in mineral identification. It may be used in two ways. As a reference tool, it will print all stored data on any of 601 different minerals. As an aid in mineral identification it will sort through all stored minerals and list only those that fit the data supplied by the user. The program requests data, sorts, and prints the names of the minerals having similar properties (up to 25).

The scan routine for refractive-index matches has a tolerance of ± .005; hardness has a tolerance of ± .51; and specific gravity has a tolerance of ± .21.

The user has the option to supply more data on the same mineral if he is not satisfied with the preliminary match Several minerals may be checked or referenced in a single run.

REFERENCES

Deer, W.A., Howie, R.A., and Zussman, Jr., *Rock Forming Minerals*, (John Wiley & Sons, Inc., New York, 1962), 5 Vol.

Larsen, E., and Berman, H., *Microscopic Determination of the Nonopaque Minerals*, (U.S. Govt. Print. Off., Wash. D.C., 1934).

Engineering & Technol.

DESCRIPTIVE TITLE	95X95 Factor Analysis with Varimax Rotation
CALLING NAME	MESA1 NUCC016
INSTALLATION NAME	Vogelback Computing Center, Northwestern University
AUTHOR(S) AND AFFILIATION(S)	Vogelback Computing Center, Northwestern University
LANGUAGE	CDC FORTRAN IV
COMPUTER	CDC 6400
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Lorrain Borman, EIN Technical Representative, Vogelback Computing Center, Northwestern University, 2129 Sheridan Road Evanston, Ill. 60201 Tel.: (312) 492-3682

FUNCTIONAL ABSTRACT

MESA1 can correlate, factor, and rotate a maximum of 95 variables. The maximum sample size that can be used with the program is 32 768 observations. Sample size should be larger than number of variables.

MESA1 is composed of three parts: a main program, an eigenvalue subroutine, and a varimax-rotation subroutine. The eigenvalues are calculated by a modified Jacobian method that closely parallels the routine given by Greenstadt in the Ralston and Wilf volume.[1] The varimax subroutine is from BIMD17 package and is based on the Kaiser algorithm.[2] The method of principal-components orthogonal rotation is used.

All of the output of the program is in the form of naturally arranged tables. The variables in the tables can be labeled by name as well as by number. The output of the program can include

—means, standard deviations, the third and fourth moments of each variable

—the standard errors of each of the above

continued

—triangular matrix of product moment correlation coefficients, with communality estimates in the main diagonal

—a table of positive eigenvalues, with percentages and cumulative percentages based upon both the total variance of the matrix and the variance accounted for by the factors to be rotated

—matrix of rotated factors

—matrix of unrotated factors. (The number of factors included in this and subsequently listed output is subject to the restrictions explained in the following section.)

The program has the following options available, which are determined by Control-Card entries.

A. Input and Communality Options

The input to the program can be either a series of N(N < 32 768) observations of M variables each ($M \leq 95$) or a MXM correlation matrix. If the input is in the form of observations, the following communality options can be computed for each variable X_i

1. All 1's
2. Squared multiple correlation of X_i with the remaining X_j, $i \neq j$
3. Absolute value of the largest correlation of X_i with X_j, $i \neq j$

If a matrix input is used, only the following options are avaiable for communality estimate for variables X_i

4. All 1's
5. Absolute value of largest correlation of X_i with X_j
6. The diagonal element r_{ii} *in the data input matrix*

No factor scores are available with options 4, 5, and 6. In these cases, Cols. 36—37 of the Problem Card *must* be blank.

B. Rotation Options

1. Rotation is optional and can be suppressed
2. If rotation is not suppressed, the following series of conditions and options determines the number of variables that will be rotated. (These conditions also determine the number of factors included in the factor matrix.)

The following conditions always hold

a. A maximum of *19 factors* can be rotated

b. Only factors with *positive eigenvalues* can be rotated

continued

209

c. A maximum of 100% of the communality of the original matrix can be accounted for by factors. (In practice, this condition means that, if the first k largest factors account for 100% of the original variance, then *no addition* factors will be calculated.)

The largest number of factors that can be rotated is the minimum of the three values determined by the above conditions. The number of factors to be rotated can be *further* restricted by the following options.

d. Specification of a *maximum number of factors*

e. Specification of the *minimum eigenvalue* to be included among rotated factors. (This specification can be a constant, or can be evaluated as the absolute value of the largest negative eigenvalue)

f. Specification of a *minimum/maximum-factor loading* for factors to be included in the rotation. (In practice, this option means that it is possible to exclude from rotation any factors that have no loadings at or above the specified value. This option should be used with caution because the program will stop calculating factors at the *first* factor that fails this maximum-factor-loading criterion. It sometimes happens that a factor will fail this min/max test when a factor with a smaller eigenvalue would not fail it.)

If any of the communality options d, e, or f are used, the minimum value, so determined, serves as the limit on the number factors to be rotated, if that value is less than the value determined by condition a, b, and c. The options d, e, and f can be suppressed by leaving the appropriate Control-Card columns blank.

REFERENCES

1. Greenstadt, J., "The Determination of the Characteristic Roots of a Matrix by the Jacobi Method," in *Mathematical Methods for Digital Computers*, Ralston and Wilf, Eds. (John Wiley & Sons, Inc., New York, 1959), Chap. 7.

2. Kaiser, H.F., "The Varimax Criterion for Analytical Rotation in Factor Analysis," Psychometrika *23*, No. 3 (Sept. 1958).

DESCRIPTIVE TITLE	Synagraphic Computer-Mapping Program
CALLING NAME	SYMAP 5.12
INSTALLATION NAME	Computing Center, The Florida State University
AUTHOR(S) AND AFFILIATION(S)	Howard T. Fisher, Northwestern Technological Institute (overall design and mathematical model) Mrs. O.G. Benson, Northwestern University Computing Center (programming)
LANGUAGE	FORTRAN IV
COMPUTER	Originally developed on IBM 709; also on IBM 360/40 and IBM 7094 at Harvard University Available in EIN through The Florida State University CDC 6400
PROGRAM AVAILABILITY	Decks, listings, and documentation presently available from The Florida State University Computing Center
CONTACT	Ray Soller, Librarian, Computing Center, The Florida State University, Tallahassee, Fla. 32306 Tel.: (904) 599-3418

FUNCTIONAL ABSTRACT

The Synagraphic Computer-Mapping Program (SYMAP) produces maps that depict spatially disposed quantitative and qualitative information. Raw data of every kind (physical, social, economic, etc.) may be related, weighted, and aggregated in a graphic format by assigning values to the coordinate locations of data points or data zones. According to the application and desired representation of data, three basic types of mapping procedure may be specified: contour, conformant, or proximal.

CONTOUR—based on the use of *contour* lines, each of which represents a value remaining constant throughout its length. The map consists of closed curves that connect all points having the same numerical

continued

value. The value at each of the different levels of contour (where a single contour level will represent a particular data value) is determined by the program, according to the scale of the map and the range of the data. Between any two contour lines, a continuous variation is assumed. Therefore, the use of contour mapping should be restricted to the representation of continuously varying information, such as topography, rainfall, or population density.

CONFORMANT—based on the *conformance* to the boundaries of a data zone. This type of mapping is best suited to data for which the representation as a continuously varying surface is inappropriate owing to the significance of physical limits or boundaries. Each predefined data zone is assigned one data value and, depending on its numeric class (range), one representative character on the map itself. Local variation within the zone boundaries will not be apparent, but will, on the average, be correct.

PROXIMAL—based on *proximity* to a data point. In appearance, this type of map is similar to the conformant map. However, point information is used here to define the data zones. Each character location on the output map is assigned the value of the *nearest* data point, using nearest-neighbor techniques. Boundaries are then assumed along the lines where these values change. Then the mapping is carried out as in the conformant type.

REFERENCES

Robertson, J. C., "The SYMAP Programme for Computer Mapping," Cartographic 108–113 (Dec. 1967); taken from a report of the Select Committee on the Ordinance Survey of Scotland.

Fisher, M., "The Laboratory for Computer Graphics," Harvard Univ. Grad. School Design Suppl. (Summer 1967).

Shepard, D., "A Two-Dimensional Interpolation Function for Irregularly Spaced Data," Harvard Univ. Grad. School Design Lab. Computer Graphics (Feb. 1968); available from the Computing Center, The Florida State University. Deals with the subject of analyzing irregularly spaced data derived from a continuous surface. A method is developed for reconstructing the surface from the sampled data. This method is the main device used to generate the maps produced by SYMAP.

000 0043

000 0043

DESCRIPTIVE TITLE	Program for Analysis of Linear Systems
CALLING NAME	PALS
INSTALLATION NAME	Vogelback Computing Center, Northwestern University
AUTHOR(S) AND AFFILIATION(S)	Dr. James Van Ness Electrical Engineering Department, Northwestern University
LANGUAGE	CDC FORTRAN IV
COMPUTER	CDC 6400
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Lorraine Borman, EIN Technical Representative, Vogelback Computing Center, Northwestern University, 2129 Sheridan Road, Evanston, Ill. 60201 Tel.: (312) 492-3682

FUNCTIONAL ABSTRACT

PALS is actually a series of subroutines combined into one large main program for (1)forming the differential equations of a linear system in matrix form, (2)finding the eigenvalues of the coefficient matrix of the differential equations, and then (3)either computing the loci of the eigenvalues as individual or groups of parameters are varied, finding the eigenvectors of the coefficient matrix, or finding the sensitivities of the eigenvalues to the parameters of the system.

Although PALS was developed initially to study the load-frequency control of a large power system, it is presently written in a general form to make it applicable to many types of systems. The system to be studied may be described in terms of a block diagram or by sets of equations. The method of forming the coefficient matrix of the differential equations from this general input is described in Ref. 1. The only restriction is that the system must be representable by a set of linear differential equations with constant coefficients. Many nonlinear systems, such as the power system, can be studied for small deviations about an operating point by linearizing the system of equations about that operating point.

continued

000 0043

The accuracy of PALS is dependent on the accuracy of the method used to find the eigenvalues and the eigenvectors of the coefficient matrix of the differential equations. Fortunately, the methods that have been chosen have proven to be very satisfactory. The QR transform developed by Francis[2] is used to find the eigenvalues and is described in Ref. 1. The eigenvectors are found by an inverse-iteration method based on a method described by Wilkinson.[3] After finding the eigenvectors, the program corrects the value found for the eigenvalues by using the Rayleigh quotient. The change in the eigenvalues is insignificant in most cases, but in the few where the accuracy of the QR transform is poor[2] the error is corrected at this step. The inverse-iteration method and the Rayleigh quotient also are described in Ref. 1.

PALS provides two methods of studying the effect of parameter changes on the system eigenvalues. The first is essentially the well-known root-loci approach. Any of the parameters in the system may be incremented through a series of steps, and the resulting change in the eigenvalues will be plotted on a complex plane. The second method finds the sensitivities of specified eigenvalues to the parameter (the derivatives of the eigenvalues to individual or groups of parameters). The sensitivity method thus gives the slope of the eigenvalue locus at a given point. It has the advantage of enabling the user to determine quickly which parameters are important in a given situation. However, the complete locus of the eigenvalue is needed to determine the effect of the parameter over any range of variation. The eigenvalue locus is found by actually incrementing the parameters, whereas the sensitivity is found by using the method described in Ref. 1.

Two versions of PALS are currently being maintained on the CDC 6400 computer. PALS1 is dimensioned for systems up to the 500th order; PALS2 is a simpler version for smaller systems, usually less than 100th order. The actual limits on each of these programs are given in Ref. 1.

REFERENCES

1. Van Ness, J.E., "PALS—A Program for Analyzing Linear Systems," Northwestern Univ. Dept. Elec. Eng. (Mar. 1969). Available from EIN for the cost of duplication and mailing.
2. Francis, J.G.F., "The QR Transformation," Computer J. *4*, 265–271 (Oct. 1961); *5*, 332–345 (Jan. 1962).
3. Wilkinson, J.H., *The Algebraic Eigenvalue Problem* (Oxford University Press, London, 1965).

DESCRIPTIVE TITLE	Synagraphic Mapping
CALLING NAME	SYMAP
INSTALLATION NAME	The Pennsylvania State University Computation Center
AUTHOR(S) AND AFFILIATION(S)	Laboratory for Computer Graphics Harvard University Graduate School of Design Adapted by Larry Rich Jeffrey Simon Larry Sinkey Department of Architectural Engineering, The Pennsylvania State University
LANGUAGE	FORTRAN IV
COMPUTER	IBM System 360/67
PROGRAM AVAILABILITY	Decks, listings, and documentation presently available
CONTACT	Dr. Daniel Bernitt, EIN Technical Representative, 105 Computer Building, The Pennsylvania State University, University Park, Pa. 16802 Tel.: (814) 865-9527

FUNCTIONAL ABSTRACT

The Synagraphic Mapping program (SYMAP) produces maps that depict spatially disposed quantitative and qualitative information. Raw data of every kind (physical, social, economic, etc.) may be related, weighted, and aggregated in a graphic format by assigning values to the coordinate locations of data points or data zones. According to the application and desired representation of data, three basic types of mapping procedure may be specified: contour, conformant, or proximal.

CONTOUR
—based on the use of *contour* lines, each of which represents a value remaining constant throughout its length. The map consists of closed curves that connect all points having the same numerical value. The value at each of the different levels of contour

continued

(where a single contour level will represent a particular data value) is determined by the program, according to the scale of the map and the range of the data. Between any two contour lines, a continuous variation is assumed. Therefore, the use of contour mapping should be restricted to the representation of continuously varying information, such as topography, rainfall, or population density.

CONFORMANT
—based on the *conformance* to the boundaries of a data zone. This type of mapping is best suited to data for which the representation as a continuously varying surface is inappropriate owing to the significance of areal limits or boundaries. Each predefined data zone is assigned one data value and, depending on its numeric class (range), one representative character on the map itself. Local variation within the zone boundaries will not be apparent, but will, on the average, be correct.

PROXIMAL
—based on *proximity* to a data point. In appearance, this type of map is similar to the conformant map. However, point information is used here to define the data zones. Each character location on the output map is assigned the value of the *nearest* data point, using nearest-neighbor techniques. Boundaries are then assumed along the lines where these values change. Then, the mapping is carried out as in the conformant type.

While the contour type of map is mose often used and the easiest to produce, the conformant and proximal maps are often more helpful in the "soft" disciplines. Output is in the form of printed pages that, if the total map size exceeds the width of the computer printed page (13 in.), may easily be glued or pasted together to form a continuous map. Also included in the output is a histogram showing frequencies for given data levels, plus several optional features.

REFERENCES

Robertson, J.C., "The SYMAP Programme for Computer Mapping," Cartographic 108–113 (Dec. 1967); taken from a report of the Select Committee on the Ordinance Survey of Scotland.

Fisher, M., "The Laboratory for Computer Graphics," Harvard Univ. Grad. School Design Suppl. (summer 1967).

continued

Shepard, D., "A Two-Dimensional Interpolation Function for Irregularly Spaced Data," Harvard Univ. Grad. School Design Lab. Computer Graphics (Feb. 1968); available from the EIN Office for the cost of duplication and mailing. Deals with the subject of analyzing irregularly spaced data derived from a continuous surface. A method is developed for reconstructing the surface from the sampled data. This method is the main device used to generate the maps produced by SYMAP.

000 0054

DESCRIPTIVE TITLE	BEEF Mathematical Subroutine
CALLING NAME	BEEFM
INSTALLATION NAME	University of Notre Dame Computation Center
AUTHOR(S) AND AFFILIATION(S)	UNIVAC Division of Sperry Rand Corporation Westinghouse Electric Corporation Baltimore Defense and Space Center Boeing Corporation
LANGUAGE	FORTRAN IV and SLEUTH II
COMPUTER	UNIVAC 1107
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Mrs. Elizabeth Hutcheson, Computer Center, University of Notre Dame, Notre Dame, Ind. 46556 Tel.: (219) 283-7784

FUNCTIONAL ABSTRACT

The BEEF mathematical library is a set of subroutines supplied by UNIVAC to "enhance FORTRAN's abilities as a scientific processor." This enrichment is in the form of subroutines for the evaluation of mathematical functions, matrix arithmetic, and other standard engineering requirements.

See the User Instructions for further information on BEEFM.

REFERENCE

UNIVAC 1107 BEEF Math Routines Manual (UP-3984), UNIVAC Division of Sperry Rand Corporation, (New York, N.Y., March 1965). Available through the UNIVAC 1107 Librarian, Systems Prog. Library Services, UNIVAC Div. of Sperry Rand, Sperry Rand Bldg., New York, N.Y. 10019.

DESCRIPTIVE TITLE	Simulation Package for University Research and Training
CALLING NAME	SPURT
INSTALLATION NAME	Vogelback Computing Center Northwestern University
AUTHOR(S) AND AFFILIATION(S)	Gustave J. Rath Department of Industrial Engineering and Management Sciences Martin Goldberg Leonard Weiner Northwestern University
LANGUAGE	CDC FORTRAN IV
COMPUTER	CDC 6400
PROGRAM AVAILABILITY	*Decks and listings presently available*
CONTACT	Lorraine Borman, EIN Technical Representative, Vogelback Computing Center, Northwestern University, 2129 Sheridan Road, Evanston, Ill. 60201 Tel.: (312) 492-3682

FUNCTIONAL ABSTRACT

SPURT is a comprehensive package of USASI Standard FORTRAN routines that are designed for use in simulation modelling. These useful routines, ranging from simple to complex, enable the average FORTRAN programmer to employ simulation techniques without having to learn the semantic and syntactic rules of a new programming language.

The SPURT package is made up of six main parts.

I. *CLOCK Generation—SPURT1*
II. *Stochastic Generators—SPURT2*
III. *Statistical Computations—SPURT3*
IV. *Analog Simulators—SPURT4*
V. *List-Processing and Queue-Manipulation—SPURT5*
VI. *Matrix and Graphical Output—SPURT6*

continued

The following is a listing and brief discussion of the various subroutines contained in each of the six SPURT parts.

CLOCK Generation—SPURT1: to implement discrete-time simulation models; to cause events to occur in the proper time sequence

The CLOCK subroutine consists basically of two lists:

> Master Time List—contains events scheduled to happen in the future
>
> Master Time Queue—contains events that could not take place at the time when they were scheduled to and, therefore, have been rescheduled; i.e., they have been blocked and are waiting in a queue.

Events can be stored on either list.

CLOCK recognizes two basic kinds of events:

> Exogenous—those that are *external* to the user's routine; these are read from Data Cards by the CLOCK
>
> Endogenous—those that are *internal* to the user's routine; these are generated dynamically and then are maintained by the CLOCK

Stochastic Generators—SPURT2: to generate samples from various probability distributions and to calculate sample values

Subroutine	*Usage*
STOGN1	Permits sampling from a discrete empirical probability distribution defined by the user
STOGN2	Enables the user to approximate a continuous distribution by means of a piecewise linear distribution
UNIFRM	Permits the user to sample real values from a uniform distribution in a defined interval
RANDIN	Provides a uniform distribution of integers in a defined interval
NORMAL	Allows the user to obtain a random sample from a normal distribution with given mean and standard deviation
NEGEXP	Permits the user to obtain a random sample from the negative exponential distribution
POISSN	Provides the user with a random sample from the Poisson distribution

continued

ERLANG	Provides a random sample from the Erlang distribution
DISCRT	Permits sampling from a step function describing a discrete cumulative distribution of integer values
LINEAR	Provides the user with a random sample from a cumulative distribution that is obtained by linear interpolation in a nonequidistant table of real values
DRAW	Provides a boolean value of TRUE or FALSE
RANPER	Generates a uniformly distributed, random permutation of the integers 1, 2, ···, M

Statistical Computations—SPURT3: to calculate statistical parameters and histograms of data arrays

Subroutine	*Usage*
STIX1 STIX2 STIX3	Three interrelated subroutines to accumulate and print out a frequency table and to produce a CalComp plot of a normalized histogram of the table
STIX4	Evaluates the mean, standard deviation, maximum value, and minimum value of an array of real numbers
STIX5	Evaluates the correlation coefficient between two arrays of real numbers
STIX6	Ranks an array of real numbers and produces the median and range of the data within the array
STIX7	Produces a statistical description of the data found in an array, including the sample size, mean, standard deviation, standard error, minimum and maximum values, range, and a printed histogram plot

Analog Simulators—SPURT4: to enable the simulation of analog-computer problems on a digital computer

Subroutine	*Usage*
ANALOG SECND	These two subroutines make it possible to obtain output similar to a hybrid computer

continued

List-Processing and Queue-Manipulation—SPURT5: lists are mXn arrays; entries in lists are mX1 arrays.

Subroutine	*Usage*
ADFIFO	Adds an entry at the bottom of the list; it can be removed only after all the elements presently on the list are gone (builds first-in last-out list)
ADLIFO	Adds an entry at the top of the list; it will be removed before any other entry presently on the list (builds last-in first-out list)
REMOVE	Removes the top (or first) entry from a list
PURGE	Destroys the contents of a list
DISPL	Prints the contents of a list

Additional subroutines in SPURT5 provide the capability to rank lists and to delete or to insert entries into lists.

Matrix and Graphical Output—SPURT6: output is facilitated through printing and graphical output

Subroutine	*Usage*
OUT	Prints out a square matrix with column and row headings
NSOUT	Prints out a nonsquare matrix with column and row headings
GRAPH	Produces two-dimensional graphs of plots, using a CalComp plotter

REFERENCES

1. Rath, G.J., "Description of the Simulation Package for University Research and Teaching," Northwestern Univ. Vogelback Computing Ctr., SPURT, Rev. E (1968; unpublished).

DESCRIPTIVE TITLE	Synagraphic Mapping Program
CALLING NAME	SYMAP
INSTALLATION NAME	Washington University Computing Facilities
AUTHOR(S) AND AFFILIATION(S)	Harvard University Graduate School of Design
LANGUAGE	FORTRAN IV
COMPUTER	IBM 360/50
PROGRAM AVAILABILITY	Proprietary; available for use but not distribution
CONTACT	Dr. Charles Drebes, Manager, Scientific Data Processing, Computing Facilities, Box 1098, Washington University, St. Louis, Mo. 63130 Tel.: (314) 863-0100 ext. 3141

FUNCTIONAL ABSTRACT

SYMAP is a computer program designed to allow city planners, geographers and others to produce low cost graphic displays of spatial patterns using standard computer line printers, by producing maps which graphically depict spatially disposed quantitative and qualitative information. It is suited to a broad range of applications, and is provided with numerous options to meet widely varying requirements. Raw data of every kind (physical, social, economic, etc.) when given to the computer may be related, manipulated, weighted, and aggregated in any manner desired. By assigning values to the coordinate locations of data points or data zones, one or more of three types of map may be produced, as specified by the user: conformant (choropleth), contour, and proximal. Potential applications are independent of the scale at which one wishes to display data. Studies (at other universities) have included a living cell, land parcels, blocks, tracts, towns, states, and continents. In each case, a common factor was the spatial distribution of a variable and a need to display the patterns associated with this distribution.

REFERENCES

Reference Manual for Synagraphic Computer Mapping—"SYMAP" (Cambridge, Mass.: Harvard Univ. Grad. Sch. of Design, Comp. Graph. Lab.).

DESCRIPTIVE TITLE	Continuous System Modeling Program
CALLING NAME	CSMP
INSTALLATION NAME	Washington University Computing Facilities
AUTHOR(S) AND AFFILIATION(S)	IBM Application Program
LANGUAGE	CSMP imbedded in FORTRAN
COMPUTER	IBM 360/50
PROGRAM AVAILABILITY	Proprietary; available for use but not for distribution.
CONTACT	Dr. C.B. Drebes, Mgr., Scientific Data Processing, Computing Facilities, Box 1098, Washington University, St. Louis, Mo. 63130 Tel.: (314) 863-0100 ext. 3141

FUNCTIONAL ABSTRACT

CSMP performs a simulation of a continuous system, obtaining solutions to problems expressed in the form of systems of differential equations or analog block diagrams. Typical applications might be a control engineer's study of the effectiveness of various control system designs, or a study of a cardiovascular system model.

CSMP provides a basic set of functional blocks with which the components of a continuous system may be represented, and it accepts application-oriented statements for defining the connections between these functional blocks. It also accepts FORTRAN IV statements which can be used to handle non-linear and time-variant problems. Input and output are facilitated by means of user-oriented control statements.

REFERENCES

System/360 Continuous System Modeling Program (360A-CX-16X) User's Manual (IBM Manual GH20-0367-3), (White Plains, N.Y.: IBM Corp. Tech. Pub. Depart., 1968).

continued

System/360 Continuous System Modeling Program (360A-CX-16X) Operator's Manual (IBM Manual H20-0368), (White Plains, N.Y.: IBM Corp. Tech. Pub. Depart., 1968).

System/360 Continuous System Modeling Program (360A-CX-16X) Application Description (IBM Manual GH20-0240-2), (White Plains, N.Y.: IBM Corp. Tech. Pub. Depart., 1968).

Humanities

DESCRIPTIVE TITLE	O'Neill Concordance Package
CALLING NAMES	CONTEXT, CONCORD
INSTALLATION NAME	Computing Center The Florida State University
AUTHOR(S) AND AFFILIATION(S)	Julia C. Lawson Computing Center The Florida State University
LANGUAGE	CDC FORTRAN IV: Main Programs CDC Assembly Language (ASCENT): Subroutines
COMPUTER	CDC 6400: Concordance Preparation IBM 1401: Card-to-Tape Input and Sorting
PROGRAM AVAILABILITY	Program decks and documentation are presently available through the CDC users group (6400) Special cases will be considered by The Florida State University Computing Center Staff
CONTACT	Dr. E.P. Miles Jr., Director Computing Center, The Florida State University, Tallahassee, Fla. 32306 Tel.: (904) 599-3418

FUNCTIONAL ABSTRACT

The O'Neill Concordance Package has been developed to prepare a concordance to the plays of Eugene O'Neill. It consists of two related programs called CONTEXT and CONCORD.

CONTEXT operates on text that has been keypunched in a relatively free format on cards. The text is edited for spacing and for certain common keypunching errors and is broken into units called *contexts*, which are usually short complete sentences or substantial clauses or phrases ending with a comma or other punctuation. The contexts, which are the input for CONCORD, are then printed for visual inspection. Occasionally, a context will contain no punctuation or may be meaningless because of the mechanical rules used. For example, the abbreviation A.F. of L. will produce three separate contexts. Such occurrences may be corrected by preparing change data for the CONCORD program.

continued

CONCORD operates on the contexts produced earlier, updating these data according to changes resulting from a visual inspection. Words that are in a user-specified dictionary are then eliminated. Each remaining word is then written on magnetic tape with its associated context, location in the text, and serial numbers.

Finally, the words are sorted into alphabetic order and merged so that all contexts for a given word are grouped together. The resulting concordance is printed on standard continuous-form paper suitable for photographic reproduction.

DESCRIPTIVE TITLE	Key Grapheme in Context
CALLING NAME	KGIC
INSTALLATION NAME	Washington University Computing Facilities
AUTHOR(S) AND AFFILIATION(S)	Joel Achtenberg Computing Facilities Washington University
LANGUAGE	PL/I
COMPUTER	S/360 under OS
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	J. Philip Miller, Computing Facilities, Washington University, St. Louis, Mo. 63130 Tel.: (314) 863-0100 ext. 4041

FUNCTIONAL ABSTRACT

The KGIC program was written to facilitate the analysis of the environmental distribution of graphic characters. It produces a KWIC-like listing of all occurrences of a given grapheme along with the graphic environment in which each instance appeared. The listing may be sorted either forward or backward from the key grapheme to facilitate inspection. Provision is made for specification by the user of special alphabets for foreign languages or for phonemic transcriptions.

If any word appears more than once in the data to be processed, only one set or records will be produced for that word. A counter will be increased and the frequency of occurrence printed in both the alphabetical listing and in the KGIC listing. The alpha listing thus will contain a complete frequency count of the corpus under consideration.

It should be noted that whenever the user can attribute phonetic or phonemic status to individual graphemes, the KGIC listing provides correspondingly significant information about phonetic and/or phonemic environments.

continued

The program produces the items listed below as output.

1. an alphabetical listing of all words processed by the KGIC program, along with their absolute and relative frequencies of occurrence,

2. the KGIC listing itself grouped by alphabetical character, with the absolute frequency of occurrence given for each unique occurrence, and the total number of occurrences and the total number of unique occurrences given for each character,

3. a summary table containing the absolute and relative frequencies of occurrence for both the total number of occurrences and for the total of unique occurences,

4. optionally, a horizontal bar graph of the relative frequencies of all occurrences of each grapheme,

5. a number of summary statistics, i.e.,
    a) total number of words processed, i.e., tokens
    b) average length of word
    c) total number of unique words, i.e., types
    d) the type/token ratio
    e) total number of characters processed
    f) total number of unique occurrences of all characters

6. a statement of all program options used in a particular run, and a complete listing of the EMICTT, defining the alphabet in use for that run.

DESCRIPTIVE TITLE	Inquirer II System for Content Analysis
CALLING NAME	I/II
INSTALLATION NAME	Washington University Computing Facilities
AUTHOR(S) AND AFFILIATION(S)	J. Philip Miller Washington University Computing Facilities
LANGUAGE	PL/I
COMPUTER	S/360 under OS
PROGRAM AVAILABILITY	Source listings and complete documentation available from Documentation Librarian, Computing Facilities, Box 1152, Washington University, St. Louis, Mo. 63130
CONTACT	J. Philip Miller, Computing Facilities, Box 1152, Washington University, St. Louis, Mo. 63130 Tel.: (314) 863-0100 ext. 4041

FUNCTIONAL ABSTRACT

The Inquirer II is a set of computer programs comparable to, but more flexible than, its predecessor, the General Inquirer, developed by Stone and his colleagues at Harvard. The original version of the General Inquirer System was designed for problems encountered in the content analysis of textual and verbal data. The General Inquirer was implemented for the IBM 7090-7094 computer along with the IBM 1401 computer. A later version of the General Inquirer was designed by Psathas and Miller to be used only with an IBM 1401 computer with an IBM 1311 disk drive. Stone originally described the General Inquirer as "a set of computer programs to (a) identify, systematically within text, instances of words and phrases that belong to categories specified by the investigator; (b) count occurrences and specify co-occurrences of these categories; (c) print and graph tabulations; (d) perform statistical tests; (e) sort and regroup sentences according to whether they contain instances of a particular category of combination of categories." The Inquirer II contains these capabilities and also

continued

allows for more elaborate analysis of the data. The I/II is able to make more elaborate contextual searches of the data and provide more options to the potential users. These options are described in detail in the *Inquirer II Programmer's Guide.*

Content analysis may be defined as a research technique which includes a systematic identification of theoretically relevant constructs in textual data. Content analysis is usually performed so that inferences can be made about the source or originator of the message, the message itself or the intended receiver of the message. The investigator communicates the constructs and the rules by which they may be identified within the corpus of text by means of dictionary. This dictionary is either one of his own construction or one which has been utilized in previous research.

REFERENCES

Miller, J. Philip, editor, *Inquirer II Programmer's Guide,* 2nd Edition, 1970, Washington University, (Available from Documentation Librarian, Computing Facilities, Box 1152, Washington University, St. Louis, Mo. for $5.00)

Stone, P.J., Dunphy, D.C., Smith, M.S., and Ogilvie, D.M. The General Inquirer: *A Computer Approach to Content Analysis*, Cambridge, Mass., The MIT Press, 1966.

Library & Inform. Sci.

DESCRIPTIVE TITLE	"Book-Type" Indexing Program
CALLING NAME	INDEXER NUCC166
INSTALLATION NAME	Vogelback Computing Center, Northwestern University
AUTHOR(S) AND AFFILIATION(S)	James Ibers, Chemistry Department, Northwestern University (original programming design) Donald Dillaman, Vogelback Computing Center, Northwestern University (CDC 6400 conversion) Lorraine Borman, Vogelback Computing Center, Northwestern University (program writeup)
LANGUAGE	CDC FORTRAN IV
COMPUTER	CDC 6400
PROGRAM AVAILABILITY	Deck and listing presently available
CONTACT	Lorraine Borman, EIN Technical Representative, Vogelback Computing Center, Northwestern University, 2129 Sheridan Road, Evanston, Ill. 60201 Tel.: ([illegible]12) 492-3682

FUNCTIONAL ABSTRACT

Purpose and General Description. INDEXER is a FORTRAN program that will produce a "book-type" index, sorting and printing under a major heading and two subheadings.

Data Input
The program sorts on 136 columns.

Format for data cards is
major heading/first subheading/second subheading/page number

An * in column 80 of the first card of a data item indicates that information is continued onto a second card.

DESCRIPTIVE TITLE	Information-Oriented Language—A Generalized Information and Retrieval System
CALLING NAME	INFOL
INSTALLATION NAME	Vogelback Computing Center, Northwestern University
AUTHOR(S) AND AFFILIATION(S)	Dr. Jacques Vallee, Control Data Corporation Robert Chalice Department of Industrial Engineering, Northwestern University (CDC 6400 version)
LANGUAGE	CDC FORTRAN IV
COMPUTER	CDC 6400
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Lorraine Borman, EIN Technical Representative, Vogelback Computing Center, Northwestern University, 2129 Sheridan Road, Evanston, Ill. 60201 Tel.: (312) 492-3682

FUNCTIONAL ABSTRACT

INFOL (Information-Oriented Language) is a generalized information storage and retrieval system that permits the user to extract selected pieces of information from a file, presented quickly and conveniently in their original form.

The INFOL language, a subset of English, was designed to be learned and used easily by those without a programming background or a knowledge of computers.

INFOL is a free-field language, thus eliminating the requirement for any part of the language to appear in specific columns of a punched card. Although punched cards are used, the concept of cards as unit records is generally avoided. The concept of an information field, implying limited word size, has been eliminated in INFOL; the user need not be concerned about how many characters to assign any piece of information. All information used as input is treated as a variable-length, but continuous, string of characters. The output may be a printed report or it may be a punched-card deck to be used as input for another computer application.

continued

INFOL includes several basic information-retrieval operations. It first allows the user to store (and later to modify) quantities of data. The system requires no input commands; files of information are established by the user with a single control word, followed by data descriptions and the data. INFOL also facilitates the selection and retrieval of information from the files according to specific requirements (retrieval criteria). The comprehensive examination (interrogation) of stored data, based on efficient retrieval criteria, is the heart of the INFOL system.

REFERENCES

INFOL for the CDC 6400 is based on INFOL as developed by the Control Data Corporation for the CDC 3600/3800 computer. Northwestern University gratefully acknowledges CDC's support in permitting it to utilize INFOL as the basis for the 6400 version and for permission to use the INFOL reference manual number 60170300 to produce its own manual, "INFOL for the CDC 6400—Information Storage and Retrieval System" (Nov. 1968). Support was also given by the U.S. Air Force Office of Scientific Research: "Man-Machine Interaction in Information Retrieval Systems."

DESCRIPTIVE TITLE	Information Retrieval System for Creating, Maintaining, Indexing, and Retrieving from Files of Textual Information
CALLING NAME	TRIAL NUCC118
INSTALLATION NAME	Vogelback Computing Center, Northwestern University
AUTHOR(S) AND AFFILIATION(S)	Donald Dillaman Lorraine Borman, Vogelback Computing Center, Northwestern University
LANGUAGE	CDC FORTRAN IV
COMPUTER	CDC 6400
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Lorraine Borman, EIN Technical Representative, Vogelback Computing Center, Northwestern University, 2129 Sheridan Road, Evanston, Ill. 60201 Tel.: (312) 492-3682

FUNCTIONAL ABSTRACT

TRIAL is an information-processing system that will perform editing, indexing, and retrieval of textual and certain types of numeric information. The system allows for the creation and maintenance of a master file (EDIT), indexing on words designated as *key* words or, alternatively, on every word in the text, excluding those common terms that are user-supplied as a "stop" word list (INDEX), and computer retrieval and printout of entries that satisfy a user search command (SEARCH). The system is designed such that any one or any combination of the above features can be achieved through one computer run with proper control cards.

The system is sufficiently flexibe to handle diverse applications in information retrieval. TRIAL has been successfully used at Northwestern in such applications as a "Selective Dissemination of Information" (SDI) system, which automatically notifies social scientists of new journal articles that appear to fit their personal interests, and as a way to select students for overseas work, by retrieving from personnel files those individuals whose backgrounds satisfied the requirements of the project. The system is especially adaptable to large masses of bibliographic data where either selective bibliographies or various forms of indexes are desired.

continued

TRIAL is actually a series of six programs—TRIAL, EDIT, SEARCH, INDEX, SORTER, and PRINT—which, respectively, function as the Executive (issuing calls to other programs as needed), the Editor (processing new input data and updating the old), the Searcher (finding instances of given key words), the Selector (retrieving and indexing key words and their contexts), the Sorter (sorting the information file), and the Printer (controlling the numerous output options).

The TRIAL file structure allows for a maximum of nine levels of record definition within an entry; an entry is considered to be all of the information for one unit of analysis. Individual applications may utilize from one to all nine designated levels of information. Any level may have new information added to it, or, if the original design of the structure needs change, new levels may be added to the existing file.

Any master file created by the EDIT program can be searched. The important features of the SEARCH program are (1) its ability to search textual material for specified words or combinations of related key words and (2) its use of the Boolean operators, AND, OR, and NOT, giving the user the ability to specify how these key words must appear in combination with each other. If the user does not retrieve relevant documents, he may rephrase his command in much the same manner as does the user of a library card catalog.

Key-word indexing by computer is an automatic method for generating alphabetized listings of key words contained in input material. Input to the INDEX overlay of the TRIAL system is any file that has been created by the EDIT program, either in the same or a previous computer run. Two major options allow for completely general and lucid indexing of key words.

REFERENCES

Janda, K., and Tetzlaff, W.H.,"TRIAL: A Computer Technique for Retrieving Information from Abstracts of Literature," Behavioral Sci., *11* (Nov. 1966).

Janda, K., and Rader, G.,"Selective Dissemination of Information: A Progress Report from Northwestern University," Amer. Behavioral Sci., *10* (Jan. 1967).

Janda, K., *Information Retrieval: Applications to Political Science* (The Bobbs-Merrill Co., Inc., New York, 1968).

continued

Tetzlaff, W.H., "TRIAL: Technique for Retrieving Information from Abstracts of Literature—A Program for the IBM 709/90/94," paper presented at the SHARE XXVI meeting, Appl. Management Sci. Proj. Committe, 24 Feb.–4 Mar. 1966, San Diego, Calif.

TRIAL: An Information Retrieval System for Creating, Maintaining, Indexing, and Retrieving from Files of Textual Information, User's Manual, Northwestern Univ. Vogelback Computing Ctr. (Oct. 1968; revised Dec. 1968).

DESCRIPTIVE TITLE	Key Grapheme in Context
CALLING NAME	KGIC
INSTALLATION NAME	Washington University Computing Facilities
AUTHOR(S) AND AFFILIATION(S)	Joel Achtenberg Computing Facilities Washington University
LANGUAGE	PL/I
COMPUTER	S/360 under OS
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	J. Philip Miller, Computing Facilities, Washington University, St. Louis, Mo. 63130 Tel.: (314) 863-0100 ext. 4041

FUNCTIONAL ABSTRACT

The KGIC program was written to facilitate the analysis of the environmental distribution of graphic characters. It produces a KWIC-like listing of all occurrences of a given grapheme along with the graphic environment in which each instance appeared. The listing may be sorted either forward or backward from the key grapheme to facilitate inspection. Provision is made for specification by the user of special alphabets for foreign languages or for phonemic transcriptions.

If any word appears more than once in the data to be processed, only one set or records will be produced for that word. A counter will be increased and the frequency of occurrence printed in both the alphabetical listing and in the KGIC listing. The alpha listing thus will contain a complete frequency count of the corpus under consideration.

It should be noted that whenever the user can attribute phonetic or phonemic status to individual graphemes, the KGIC listing provides correspondingly significant information about phonetic and/or phonemic environments.

continued

The program produces the items listed below as output.

1. an alphabetical listing of all words processed by the KGIC program, along with their absolute and relative frequencies of occurrence,

2. the KGIC listing itself grouped by alphabetical character, with the absolute frequency of occurrence given for each unique occurrence, and the total number of occurrences and the total number of unique occurrences given for each character,

3. a summary table containing the absolute and relative frequencies of occurrence for both the total number of occurrences and for the total of unique occurences,

4. optionally, a horizontal bar graph of the relative frequencies of all occurrences of each grapheme,

5. a number of summary statistics, i.e.,

   a) total number of words processed, i.e., tokens
   b) average length of word
   c) total number of unique words, i.e., types
   d) the type/token ratio
   e) total number of characters processed
   f) total number of unique occurrences of all characters

6. a statement of all program options used in a particular run, and a complete listing of the EMICTT, defining the alphabet in use for that run.

DESCRIPTIVE TITLE	BEEF Data Processing Subroutines
CALLING NAME	BEEFDP
INSTALLATION NAME	The University of Notre Dame Computing Center
AUTHOR(S) AND AFFILIATION(S)	Westinghouse Electric Corporation UNIVAC Division of Sperry Rand Corporation
LANGUAGE	SLEUTH II/FORTRAN IV
COMPUTER	UNIVAC 1107
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Elizabeth Hutcheson, EIN Technical Representative, Computing Center, University of Notre Dame, Notre Dame, Ind. 46556 Tel.: (219) 283-7784

FUNCTIONAL ABSTRACT

The BEEF Data Processing library is a set of subroutines supplied by UNIVAC to "enhance FORTRAN's abilities as a commercial processor." This enrichment is in the form of subroutines for whole word data movement, character and field movement, supplementary formatting, decision making with FORTRAN arrays, data conversion, report generation and control, I/O control, sorting, and compatability with EAM/EDP (Electronic Accounting Machine/Electronic Data Processing) equipment.

REFERENCE

UNIVAC 1107 BEEF Data Processing Manual, (UP-3986), (UNIVAC Division of Sperry Rand Corp., N.Y.). Copies of this manual are available through the local UNIVAC representative or through the EIN Office at the cost of reproduction and mailing.

000 0058

DESCRIPTIVE TITLE	Inquirer II System for Content Analysis
CALLING NAME	I/II
INSTALLATION NAME	Washington University Computing Facilities
AUTHOR(S) AND AFFILIATION(S)	J. Philip Miller Washington University Computing Facilities
LANGUAGE	PL/I
COMPUTER	S/360 under OS
PROGRAM AVAILABILITY	Source listings and complete documentation available from Documentation Librarian, Computing Facilities, Box 1152, Washington University, St. Louis, Mo. 63130
CONTACT	J. Philip Miller, Computing Facilities, Box 1152, Washington University, St. Louis, Mo. 63130 Tel.: (314) 863-0100 ext. 4041

FUNCTIONAL ABSTRACT

The Inquirer II is a set of computer programs comparable to, but more flexible than, its predecessor, the General Inquirer, developed by Stone and his colleagues at Harvard. The original version of the General Inquirer System was designed for problems encountered in the content analysis of textual and verbal data. The General Inquirer was implemented for the IBM 7090-7094 computer along with the IBM 1401 computer. A later version of the General Inquirer was designed by Psathas and Miller to be used only with an IBM 1401 computer with an IBM 1311 disk drive. Stone originally described the General Inquirer as "a set of computer programs to (a) identify, systematically within text, instances of words and phrases that belong to categories specified by the investigator; (b) count occurrences and specify co-occurrences of these categories; (c) print and graph tabulations; (d) perform statistical tests; (e) sort and regroup sentences according to whether they contain instances of a particular category of combination of categories." The Inquirer II contains these capabilities and also

continued

000 0058

allows for more elaborate analysis of the data. The I/II is able to make more elaborate contextual searches of the data and provide more options to the potential users. These options are described in detail in the *Inquirer II Programmer's Guide*.

Content analysis may be defined as a research technique which includes a systematic identification of theoretically relevant constructs in textual data. Content analysis is usually performed so that inferences can be made about the source or originator of the message, the message itself or the intended receiver of the message. The investigator communicates the constructs and the rules by which they may be identified within the corpus of text by means of dictionary. This dictionary is either one of his own construction or one which has been utilized in previous research.

REFERENCES

Miller, J. Philip, editor, *Inquirer II Programmer's Guide*, 2nd Edition, 1970, Washington University, (Available from Documentation Librarian, Computing Facilities, Box 1152, Washington University, St. Louis, Mo. for $5.00)

Stone, P.J., Dunphy, D.C., Smith, M.S., and Ogilvie, D.M. The General Inquirer: *A Computer Approach to Content Analysis*, Cambridge, Mass., The MIT Press, 1966.

DESCRIPTIVE TITLE	Basic Information Retrieval System
CALLING NAME	BIRS
INSTALLATION NAME	Michigan State University Information Systems Laboratory
AUTHOR(S) AND AFFILIATION(S)	John F. Vinsonhaler, Ph.D. John M. Hafterson Stuart W. Thomas, Jr. Michigan State University
LANGUAGE	USASI Full FORTRAN
COMPUTER	CDC 3600, CDC 6500, CDC 6600 IBM 360 G-level GE 600 Series, others
PROGRAM AVAILABILITY	Decks and listings are currently available at cost from Michigan State University for non-profit institutions. The program is distributed to profit-making institutions by Hygain Technologies, 65 Whitney Street, Westport, Conn. Maintenance for all users is provided by Hygain Technologies.
CONTACT	Dr. John F. Vinsonhaler, Director, Information Systems Laboratory, 309 Computer Center, Michigan State University, East Lansing, Mich. 48823 Tel.: (517) 353-7284

FUNCTIONAL ABSTRACT

BIRS is a general purpose system of programs for the behavioral sciences and education. Essentially, BIRS is a set of fundamental program modules designed to allow scholars and scientists to use their own locally based computer to construct and maintain a variety of information systems. Search, maintenance, and index creation are performed automatically. Thus, BIRS may be viewed as a set of essential tools; the research worker may use these tools to construct the type of information system which best meets his immediate needs.

continued

It is not economically practical for individual educators and social scientists to develop their own special purpose programming system; they must share the costs of system developments by exchanging programs. General purpose systems like BIRS (with machine, data, and application independent programming) are ideally suited for free exchange among computer users.

REFERENCES

Vinsonhaler, J.F., Ed., *Technical Manual for the Basic Indexing and Retrieval System, BIRS 2.0*, Michigan State University, East Lansing, Mich., 1968.

Vinsonhaler, J.F., and Hafterson, J.M., Eds., *Technical Manual for the Basic Indexing and Retrieval System, Appendix I, BIRS 2.5*, Michigan State University, East Lansing, Mich., 1969.

Thomas, Jr., S.W., and Vinsonhaler, J.F., Eds., *Technical Manual for the Basic Information Analysis System, BIRS Technical Manual, Appendix II, BIAS 1.0*, Michigan State University, East Lansing, Mich., 1969.

Documentation is available at cost from the Information Systems Laboratory and from Hygain Technologies.

DESCRIPTIVE TITLE	TEXT360: A System for Producing Manuals
CALLING NAME	TEXT360
INSTALLATION NAME	Washington University Computing Facilities
AUTHOR(S) AND AFFILIATION(S)	S.L. Reed, International Business Machines
LANGUAGE	PL/I and Assembly Language
COMPUTER	IBM/360 under OS
PROGRAM AVAILABILITY	Proprietary; available for use but not for distribution
CONTACT	J. Philip Miller, Computing Facilities, Box 1152, Washington University, St. Louis, Mo. 63130 Tel.: (314) 863-0100 ext. 3141

FUNCTIONAL ABSTRACT

TEXT360 is a text-processing system with data-entry, data-updating, and page-formatting capabilities. The system, which runs under Operating System/360, consists of a main processor and several peripheral programs. Input to the system is free-form and is produced on the IBM 029 Card Punch. Output is camera-ready and is produced on the IBM 1403 Printer. The formatting capabilities of TEXT360 permit the insertion, deletion, and replacement of characters, words, lines, and groups of lines. In addition, blocks of text can be moved from one part of a document to another. One-column and two-column page format can be produced. Routine functions include hyphenation, line justification, column heading, and indentations. More complex functions include the generation of horizontal and vertical ruling for tables and figures. The program also allows the user to specify that related material (e.g., a table) is to be kept together, i.e., not split between columns or pages.

The four phases of the TEXT360 Formatting Processor, the TEXT360 Spelling Dictionary Update Program, the TEXT90-to-TEXT360 Master File Conversion Program and the Print/Punch Utility Program are

continued

written in the PL/1 Language, supplemented by four small assembler-language routines used for character-set mapping. The TEXT360 Prescan and Peripheral Print Programs are written in System/360 assembler language.

TEXT360 is essentially the program used by IBM for the production of the SRL manuals.

REFERENCES

Reed, S.L., *TEXT360*, Contributed Program Library Documentation, Program 360D-29.4.001 (White Plains, N. Y.: IBM, 1967).

DESCRIPTIVE TITLE	Information Retrieval Services
CALLING NAME	
INSTALLATION NAME	University of Georgia Computer Center
AUTHOR(S) AND AFFILIATION(S)	Information Science Group University of Georgia
LANGUAGE	Assembler Language & PL/1
COMPUTER	IBM 360/65
PROGRAM AVAILABILITY	Proprietary; available for use but not for distribution
CONTACT	Miss Margaret Park, Supervisory Inform. Scientist, Computer Center, The Univ. of Ga., Athens, Ga. 30601 Tel.: (404) 542-3741

FUNCTIONAL ABSTRACT

In any endeavor, scientists need to keep constantly abreast of activities in their field of interest, to be on the lookout for new ideas, and to maintain a library of useful references. The proliferation of new scientific knowledge is rapidly outpacing the capabilities of conventional information-handling and publishing techniques. Scientists are now turning to computer-based methods to help speed and channel the flow of information on a timely basis.

The Computer Center at the University of Georgia is actively engaged in establishing an Information Center. Mechanized data bases from several scientific organizations are presently available and in use at the Center. Subject areas currently represented are biology, biochemistry, nuclear science, and chemistry, including structural data files for chemical compounds. Other tape services in fields such as medicine, engineering, physics, geology, etc., will be added as interst is expressed in these subject areas.

The Computer Center's Information Sciences Unit offers assistance in creating search profiles and current awareness and retrospective searches of the scientific literature.

continued

The staff of the Computer Center includes highly trained professionals in practically every subject discipline. In addition to chemists and a microbiologist in the Information Sciences Unit, the Center also has full time staff with specialties in physics, engineering, statistics, biology, and forestry, with consulting staff available in other disciplines. All staff are also familiar with computer systems and their applications.

The Computer Center is constantly seeking better ways to satisfy the information requirements of the scientific community and we will greatly appreciate any suggestions made in this direction. New services will be added as rapidly as possible to meet newly identified needs and uses.

Persons desiring to use the Information Retrieval Services are directed to the contact person.

DESCRIPTIVE TITLE	FORTRAN Program for Computer Based Serials Holdings Management
CALLING NAMES	UPDATE, HOLD, PUBLSH, LANSUB
INSTALLATION NAME	Indiana University-Purdue University at Indianapolis Research Computation Center
AUTHOR(S) AND AFFILIATION(S)	Mrs. Alma Connell School of Medicine Library Indiana University Mrs. J. Mueller School of Medicine Library Indiana University Mrs. Judy Silence Research Computation Center, IUPUI
LANGUAGE	IBM FORTRAN IV (IBSYS)
COMPUTER	IBM 7040
PROGRAM AVAILABILITY	Deck and listing presently available
CONTACT	Dr. David A. Neal, EIN Tech. Rep., Research Comp. Ctr., Indiana Univ.—Purdue Univ. at Indianapolis, 1100 West Michigan St., Indianapolis, Ind. 46202 Tel.: (317) 639-7813

FUNCTIONAL ABSTRACT

This system is a long range project concerning the holdings of serials in the School of Medicine Library. Three overall applications are (1) updating the file to maintain a current holding list, (2) publication of selected lists using search techniques, and (3) annual publication of complete holdings list.

Several programs have been prepared for these applications. Detailed descriptions are given in Ref. 1. A general description of each program follows.

System Update Program 1

This program produces a readable list for visual checking of every item on the cards. When a set of cards has been keypunched,

continued

they are run with the System Update Program 1 using the card reader as the input unit. The output is checked and the cards are corrected. The cards are then added to the master file with the UPDATE (number 2) program.

System Update Program 2

This program builds a master tape which may be used as data for the programs that follow. UPDATE will also be used for maintaining the current file, i.e., inserting new cards as they are punched and replacing cards when new information is needed.

HOLD

HOLD can be run with the master file to produce a listing of the information on each card on that file. HOLD also provides a summary table giving the number of titles for specified categories.

PUBLSH

PUBLSH provides thirteen options for listing holdings for publication. These are (1) full list with holdings, (2 and 3) nursing list with or without holdings, (4 and 5) currently received list with or without holdings, (6 and 7) Index Medicus list with or without holdings, (8 and 9) Indexes, Abstracts and Bibliographies list with or without holdings, (10 and 11) International Nursing Index with or without holdings, and (12 and 13) microform with or without holdings. Multiple copies of output from this program are available.

LANSUB

LANSUB provides four general options for title listings. These are (1) language only, (2) one subject, (3) two subjects listed according to "and" or "or" logic, and (4) a language and a subject cross-indexed.

Listings of complete holdings or of currently received serials only can be obtained under the four options above. On complete holdings lists, currently received titles are indicated by a plus (+) to the left of the title. Multiple copies of output from this program are available.

REFERENCE

1. Indiana University Medical Center Research Computation Center Library, *Program Description* (Indianapolis, Ind: Indiana Univ., 1966, Rev. 1968). Available from the EIN Office at the cost of reproduction and mailing.

Life Sciences

DESCRIPTIVE TITLE	Automatic Interaction Detector
CALLING NAME	AID NUCC113
INSTALLATION NAME	Vogelback Computing Center, Northwestern University
AUTHOR(S) AND AFFILIATION(S)	J. Sonquist The University of Michigan
LANGUAGE	CDC FORTRAN IV
COMPUTER	CDC 6400
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Lorraine Borman, EIN Technical Representative, Vogelback Computing Center, Northwestern University, 2129 Sheridan Road Evanston, Ill. 60201 Tel.: (312) 492-3682

FUNCTIONAL ABSTRACT

AID was originally programmed for the IBM 7090 in the MAD language at The University of Michigan. A detailed discussion of the theory, methods, and control parameters of the program are contained in Ref. 1. Since 1964, two parameters have been added to the main-parameter card.

AID is focused on a particular kind of data-analysis problem, characteristic of many social-science research situations, in which the purpose of the analysis involves more than the reporting of descriptive statistics but may not necessarily involve the exact testing of specific hypotheses. In this type of situations the problem is often one of determining which of the variables, for which data have been collected, are related to the phenonmenon in question, under what conditions, and through what intervening processes, with appropriate controls for spuriousness.

AID is useful in studying the interrelationships among a set of up to 37 variables. Regarding one of the variables as a dependent variable, the analysis employs a nonsymmetrical branching process, based on variance-analysis techniques, to subdivide the sample into a series of subgroups that maximizes one's ability to predict values of the dependent variable. Linearity and additivity

continued

assumptions inherent in conventional multiple-regression techniques are not required. AID will handle variables that are only nominal scales, i.e., mere classifications.

REFERENCES

1. Morgan, J.N., and Sonquist, J.A., "The Detection of Interaction Effects: A Report on a Computer Program for the Selection of Optimal Combinations of Explanatory Variables," Univ. Mich. Inst. Soc. Res. Survey Res. Ctr. Monograph No. 35 (1964); copies may be obtained from the EIN office for the cost of duplication and mailing.

DESCRIPTIVE TITLE	Bio-Medical Multivariate Statistical Programs
CALLING NAME	BMD
INSTALLATION NAME	University of Notre Dame Computing Center
AUTHOR(S) AND AFFILIATION(S)	School of Medicine University of California, Los Angeles UNIVAC Division of Sperry Rand Corporation College of Business Administration and Computing Center University of Notre Dame
LANGUAGE	FORTRAN IV
COMPUTER	UNIVAC 1107
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Elizabeth Hutcheson, EIN Technical Representative, Computing Center, University of Notre Dame, Notre Dame, Ind. 46556 Tel.: (219) 283-7784

FUNCTIONAL ABSTRACT

The BMD system is a package of computer programs designed to do both basic data processing and the subsequent statistical analysis. The programs have been prepared in an easy-to-use parametric form so that the researcher may adapt them to a wide variety of statistical problems. For further details of the package, see the User Instructions. The BMD is available at the University of Notre Dame in its 1967 edition.

REFERENCES

Dixon, W.J. (Ed.), *BMD: Biomedical Computer Programs*, (Univ. of Calif. Press, Berkeley, 1967, 1970).

Univ. of Notre Dame Computing Center, *BMD for the UNIVAC 1107*, (Rough draft, 1967 ed. of BMD). Available from Univ. of Notre Dame Computing Center, Notre Dame, Ind.

DESCRIPTIVE TITLE	Analysis of Change-over Experiments
CALLING NAME	ZFE-03, ZFE-04
INSTALLATION NAME	Office of Data Analysis Research Educational Testing Service (ETS)
AUTHOR(S) AND AFFILIATION(S)	Procedure: G. Beall formerly of Gillette Razor Co. Program: V. Halfmann formerly of ETS
LANGUAGE	FORTRAN
COMPUTER	IBM 360/65
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Mr. Ernest Anastasio, Off. of Data Anal. Research, Educational Testing Service, Rosedale Road, Princeton, N.J. 08540 Tel.: (609) 921-9000 ext. 2552

FUNCTIONAL ABSTRACT

The class of designs covered by this program is latin squares or Youden rectangles (incomplete latin squares). These may be repeated fully or in part. The design may be defective, i.e., certain whole rows may be missing, but no allowance has been made for missing cells, i.e., single observations.

The purpose of this program is primarily to analyze the results for effect of treatment with allowance for carry-over of preceding treatment. There is also direct testing of the significance of carry-over. There is included parallel estimation and testing of significance without allowance for carry-over. The program is self-contained and does not require any external subroutines, such as might be presumed to exist in one form or another at computation centers.

The program, as presently stored, allows analysis for designs up to 40 rows (or blocks). This is limited by the dimensions of the data matrices. It could be readily enough changed to 500 or 1000 if such an experiment were involved. The analysis has been contrived so that such change does not increase the size of the matrix involved in equation solving.

continued

Explanation

The basic equations are,

1. $y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_k + \varepsilon_{ijk}$

where y_{ijk} is an observation assumed built of a general level μ, effect of the i^{th} row or individual α_i, the j^{th} period or column β_j, the k^{th} treatment γ_k and extraneous variability ε_{ijk}. This equation obtains for the first column or period when there has been no conditioning period. For the following periods,

2. $y_{ijk\ell} = \mu + \alpha_i + \beta_j + \gamma_k + \delta_\ell + \varepsilon_{ijk\ell}$

where the effect of the ℓ^{th} treatment in the preceding period is δ_ℓ. For a conditioned experiment Equ. (2) obtains in all columns.

The data actually considered are the differences within rows such as,

$$y_{ijk\ell} - y_{i'j'k'\ell'} = \beta_j - \beta_{j'} + \gamma_k - \gamma_{k'} + \delta_\ell - \delta_{\ell'}$$
$$(j' \neq j\ ,\ k' \neq k\ ,\ \ell' \neq \ell)\ .$$

These differences are then set forth in a matrix. Thus for an unconditioned latin square for which the first line is,

Design:	(1)	(2)	(4)	(3)
Result:	4	5	7	6

we may consider the two differences,

$$y_{111} - y_{1221} = \beta_1 - \beta_2 + \gamma_1 - \gamma_2 - \delta_1 + \varepsilon_{111} - \varepsilon_{1221}$$
$$y_{1221} - y_{1342} = \beta_2 - \beta_3 + \gamma_2 - \gamma_4 + \delta_1 - \delta_2 + \varepsilon_{1221} - \varepsilon_{1342}$$

which results in two lines of the matrix as follows:

β_1	β_2	β_3	β_4	γ_1	γ_2	γ_3	γ_4	δ_1	δ_2	δ_3	δ_4	Result
+1	−1			+1	−1			−1				−1
	+1	−1			+1		−1	+1	−1			−2

Least-squares equations are in the same form. For instance, to get the equation associated with β_2, each line is multiplied by its content in the β_2 column and the product accumulated over all columns. For each set of effects (β, γ or δ), the last equation is replaced by a condition equation,

continued

$$\sum_j \hat{\beta}_j = \sum \hat{\gamma}_k = \sum_\ell \hat{\delta}_\ell = 0 \ .$$

The analysis without carry-over (SANS DELTA) is gotten by replacing, temporarily, all equations appropriate to δ by $\hat{\delta}_\ell = 0$.

The analysis without treatment (SANS GAMMA) is obtained by the temporary replacement, $\hat{\gamma}_k = 0$.

The residual variability is gotten in several steps. First the variability residual on β , γ and δ is

$$3. \quad \sum_{ijk} y^2_{ijk} + \sum_{ijk\ell} y^2_{ijk\ell} - \sum_j \hat{\beta}_j \Big(\sum_{ik} y_{ijk} + \sum_{ik\ell} y_{ijk\ell}\Big)$$

$$- \sum_k \hat{\gamma}_k \Big(\sum_{ij} y_{ijk} + \sum_{ij\ell} y_{ijk\ell}\Big) - \sum_\ell \hat{\delta}_\ell \sum_{ijk} y_{ijk\ell} \ .$$

Secondly, an estimate of $\hat{\mu}$ is made by finding from Equ. (1) or (2) the mean of the values

$$y'_{ijk} = y_{ijk} - \hat{\beta}_j - \hat{\gamma}_k$$

$$y'_{ijk\ell} = y_{ijk\ell} - \hat{\beta}_j - \hat{\gamma}_k - \hat{\delta}_\ell \quad .$$

Thirdly, estimates of $\hat{\alpha}_i$ are made by finding the mean of the values

$$y''_{ijk} = y_{ijk} - \hat{\mu}$$

$$y''_{ijk\ell} = y'_{ijk\ell} - \hat{\mu} \ .$$

The residual variability as from equation (3) is then further and finally reduced by

$$\hat{\mu} \sum_{ijk\ell} y_{ijk\ell} + \sum_i \alpha_i \sum_{jk\ell} y_{ijk\ell} \ .$$

Residual variability on the effects of $\hat{\mu}$, rows and columns only (SANS DELTA & SANS GAMMA) is gotten by the formula familiar in analysis of variance.

continued

259

The test of significance for treatments without allowance for carry-over is based on residual variability SANS DELTA less residual variability SANS DELTA & SANS GAMMA. The test with allowance for carry-over is from residual variability on FULL MATRIX less that on SANS GAMMA. The test for carry-over is from residual variability on FULL MATRIX less that on SANS DELTA.

It need only be added that there is incorporated a test on whether the situation is underdetermined. The program counts the number of different row patterns, multiples this number by the number of columns, and checks whether the result exceeds the number of independent parameters to be estimated. In the case of underdetermination, it refuses to analyze. A second type of refusal arises if the simultaneous equations prove insoluble, which may arise if the design is redundant. Finally, if there is no residual freedom, the program will estimate the parameters but declare F = 0.

DESCRIPTIVE TITLE	Information Retrieval Services
CALLING NAME	
INSTALLATION NAME	University of Georgia Computer Center
AUTHOR(S) AND AFFILIATION(S)	Information Science Group University of Georgia
LANGUAGE	Assembler Language & PL/1
COMPUTER	IBM 360/65
PROGRAM AVAILABILITY	Proprietary; available for use but not for distribution
CONTACT	Miss Margaret Park, Supervisory Inform. Scientist, Computer Center, The Univ. of Ga., Athens, Ga. 30601 Tel.: (404) 542-3741

FUNCTIONAL ABSTRACT

In any endeavor, scientists need to keep constantly abreast of activities in their field of interest, to be on the lookout for new ideas, and to maintain a library of useful references. The proliferation of new scientific knowledge is rapidly outpacing the capabilities of conventional information-handling and publishing techniques. Scientists are now turning to computer-based methods to help speed and channel the flow of information on a timely basis.

The Computer Center at the University of Georgia is actively engaged in establishing an Information Center. Mechanized data bases from several scientific organizations are presently available and in use at the Center. Subject areas currently represented are biology, biochemistry, nuclear science, and chemistry, including structural data files for chemical compounds. Other tape services in fields such as medicine, engineering, physics, geology, etc., will be added as interst is expressed in these subject areas.

The Computer Center's Information Sciences Unit offers assistance in creating search profiles and current awareness and retrospective searches of the scientific literature.

continued

The staff of the Computer Center includes highly trained professionals in practically every subject discipline. In addition to chemists and a microbiologist in the Information Sciences Unit, the Center also has full time staff with specialties in physics, engineering, statistics, biology, and forestry, with consulting staff available in other disciplines. All staff are also familiar with computer systems and their applications.

The Computer Center is constantly seeking better ways to satisfy the information requirements of the scientific community and we will greatly appreciate any suggestions made in this direction. New services will be added as rapidly as possible to meet newly identified needs and uses.

Persons desiring to use the Information Retrieval Services are directed to the contact person.

DESCRIPTIVE TITLE	Biomedical Computer Programs
CALLING NAME	BMD
INSTALLATION NAME	Washington University Computing Facilities
AUTHOR(S) AND AFFILIATION(S)	Health Sciences Computing Facility University of California, Los Angeles
LANGUAGE	FORTRAN IV
COMPUTER	IBM 360/50
PROGRAM AVAILABILITY	Proprietary; available for use at Washington University. Available for distribution from authors.
CONTACT	Dr. C. B. Drebes, Mgr., Scientific Data Processing, Computing Facilities, Box 1098, Washington University, St. Louis, Mo. 63130 Tel.: (314) 863-0100 ext. 3141

FUNCTIONAL ABSTRACT

The BMD package contains a variety of statistical programs in the following areas.

Description and Tabulation

Multivariate Analysis
a. Factor Analysis
b. Discriminant Analysis
c. Canonical Analysis

Regression Analysis
a. Linear
b. Polynomial
c. Asymptotic

Analysis of Variance and Covariance

Time Series Analysis

Special Programs
a. Life Table and Survival Rate
b. Contingency Table Analysis
c. Biological Assay
d. Guttman Scaling

continued

REFERENCES

Dixon, W.J., Ed., *BMD: Biomedical Computer Programs* (Berkeley: Univ. of Cal. Press, 1968).

Dixon, W.J., Ed., *BMD: Biomedical Computer Programs, X-series Supplement* (Berkeley: Univ. of Cal. Press, 1969).

DESCRIPTIVE TITLE	FORTRAN Program for Computer Based Serials Holdings Management
CALLING NAMES	UPDATE, HOLD, PUBLSH, LANSUB
INSTALLATION NAME	Indiana University-Purdue University at Indianapolis Research Computation Center
AUTHOR(S) AND AFFILIATION(S)	Mrs. Alma Connell School of Medicine Library Indiana University Mrs. J. Mueller School of Medicine Library Indiana University Mrs. Judy Silence Research Computation Center, IUPUI
LANGUAGE	IBM FORTRAN IV (IBSYS)
COMPUTER	IBM 7040
PROGRAM AVAILABILITY	Deck and listing presently available
CONTACT	Dr. David A. Neal, EIN Tech. Rep., Research Comp. Ctr., Indiana Univ.—Purdue Univ. at Indianapolis, 1100 West Michigan St., Indianapolis, Ind. 46202 Tel.: (317) 639-7813

FUNCTIONAL ABSTRACT

This system is a long range project concerning the holdings of serials in the School of Medicine Library. Three overall applications are (1) updating the file to maintain a current holding list, (2) publication of selected lists using search techniques, and (3) annual publication of complete holdings list.

Several programs have been prepared for these applications. Detailed descriptions are given in Ref. 1. A general description of each program follows.

System Update Program 1

This program produces a readable list for visual checking of every item on the cards. When a set of cards has been keypunched,

continued

they are run with the System Update Program 1 using the card reader as the input unit. The output is checked and the cards are corrected. The cards are then added to the master file with the UPDATE (number 2) program.

System Update Program 2

This program builds a master tape which may be used as data for the programs that follow. UPDATE will also be used for maintaining the current file, i.e., inserting new cards as they are punched and replacing cards when new information is needed.

HOLD

HOLD can be run with the master file to produce a listing of the information on each card on that file. HOLD also provides a summary table giving the number of titles for specified categories.

PUBLSH

PUBLSH provides thirteen options for listing holdings for publication. These are (1) full list with holdings, (2 and 3) nursing list with or without holdings, (4 and 5) currently received list with or without holdings, (6 and 7) Index Medicus list with or without holdings, (8 and 9) Indexes, Abstracts and Bibliographies list with or without holdings, (10 and 11) International Nursing Index with or without holdings, and (12 and 13) microform with or without holdings. Multiple copies of output from this program are available.

LANSUB

LANSUB provides four general options for title listings. These are (1) language only, (2) one subject, (3) two subjects listed according to "and" or "or" logic, and (4) a language and a subject cross—indexed.

Listings of complete holdings or of currently received serials only can be obtained under the four options above. On complete holdings lists, currently received titles are indicated by a plus (+) to the left of the title. Multiple copies of output from this program are available.

REFERENCE

1. Indiana University Medical Center Research Computation Center Library, *Program Description* (Indianapolis, Ind: Indiana Univ., 1966, Rev. 1968). Available from the EIN Office at the cost of reproduction and mailing.

Mathematics

DESCRIPTIVE TITLE	Eigensystems of Nonsymmetric, Real Square Matrices
CALLING NAME	EIGSYS NUCC126
INSTALLATION NAME	Vogelback Computing Center, Northwestern University
AUTHOR(S) AND AFFILIATION(S)	James Van Ness Department of Electrical Engineering, Northwestern University
LANGUAGE	CDC FORTRAN IV
COMPUTER	CDC 6400
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Lorraine Borman, EIN Technical Representative, Vogelback Computing Center, Northwestern University, 2129 Sheridan Road, Evanston, Ill. 60201. Tel.: (312) 492-3682

FUNCTIONAL ABSTRACT

EIGSYS produces all eigenvalues and eigenvectors (both real and complex) of either symmetric or nonsymmetric square matrices having real elements. EIGSYS calls various subroutines to calculate the eigenvalues, the determinant and the eigenvectors. The trace and determinant of input matrix A is compared with the sum and product, respectively, of the E values. The righthand eigenvectors X(I) and the lefthand eigenvectors V(J) of A are printed along with their respective residual vectors (A*X(I) - LAMBDA(I)*X(I)) and V(J)*A-V(J)*LAMBDA(J)). The squares of the residual vectors are printed for convenience in checking the accuracy of the eigensystem. After finding each E vector, EIGSYS then uses the Rayleigh quotient[1] scheme to correct the original E value, which in turn produces a corrected E vector along with new residuals. A second correction is applied that produces a further refinement of the eigenvalue only. The sum and product of the corrected E values are printed at the end of the output for each matrix. EIGSYS repeats this computation for each matrix in the data-input stream.

REFERENCES

1. Martin, R.S., and Wilkinson, J.H., "Solution of Symmetric and Asymmetric Band Equations and the Calculation of Eigenvectors of Band Matrices," Numer. Math. 8, 279–301 (1959).

DESCRIPTIVE TITLE	Linear Programming
CALLING NAME	LINPROG NUCC120
INSTALLATION NAME	Vogelback Computing Center, Northwestern University
AUTHOR(S) AND AFFILIATION(S)	School of Business, Northwestern University
LANGUAGE	CDC FORTRAN IV
COMPUTER	CDC 6400
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Lorraine Borman, EIN Technical Representative, Vogelback Computing Center, Northwestern University, 2129 Sheridan Road Evanston, Ill. 60201 Tel.: (312) 492-3682

FUNCTIONAL ABSTRACT

LINPROG will solve any linear-programming problem with a maximum of 125 variables and 75 constraint equations. It utilizes the original (two-phase) simplex method for solving linear-programming problems and uses a *minimizing* objective function. To solve a maximizing problem, the signs on the objective function coefficients must be reversed before usage.

DESCRIPTIVE TITLE	95X95 Factor Analysis with Varimax Rotation
CALLING NAME	MESA1 NUCC016
INSTALLATION NAME	Vogelback Computing Center, Northwestern University
AUTHOR(S) AND AFFILIATION(S)	Vogelback Computing Center, Northwestern University
LANGUAGE	CDC FORTRAN IV
COMPUTER	CDC 6400
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Lorrain Borman, EIN Technical Representative, Vogelback Computing Center, Northwestern University, 2129 Sheridan Road Evanston, Ill. 60201 Tel.: (312) 492-3682

FUNCTIONAL ABSTRACT

MESA1 can correlate, factor, and rotate a maximum of 95 variables. The maximum sample size that can be used with the program is 32 768 observations. Sample size should be larger than number of variables.

MESA1 is composed of three parts: a main program, an eigenvalue subroutine, and a varimax-rotation subroutine. The eigenvalues are calculated by a modified Jacobian method that closely parallels the routine given by Greenstadt in the Ralston and Wilf volume.[1] The varimax subroutine is from BIMD17 package and is based on the Kaiser algorithm.[2] The method of principal-components orthogonal rotation is used.

All of the output of the program is in the form of naturally arranged tables. The variables in the tables can be labeled by name as well as by number. The output of the program can include

—means, standard deviations, the third and fourth moments of each variable

—the standard errors of each of the above

continued

—triangular matrix of product moment correlation coefficients, with communality estimates in the main diagonal

—a table of positive eigenvalues, with percentages and cumulative percentages based upon both the total variance of the matrix and the variance accounted for by the factors to be rotated

—matrix of rotated factors

—matrix of unrotated factors. (The number of factors included in this and subsequently listed output is subject to the restrictions explained in the following section.)

The program has the following options available, which are determined by Control-Card entries.

A. Input and Communality Options

The input to the program can be either a series of $N (N < 32\ 768)$ observations of M variables each ($M < 95$) or a MXM correlation matrix. If the input is in the form of observations, the following communality options can be computed for each variable X_i

1. All 1's
2. Squared multiple correlation of X_i with the remaining X_j, $i \neq j$
3. Absolute value of the largest correlation of X_i with X_j, $i \neq j$

If a matrix input is used, only the following options are avaiable for communality estimate for variables X_i

4. All 1's
5. Absolute value of largest correlation of X_i with X_j
6. The diagonal element r_{ii} *in the data input matrix*

No factor scores are available with options 4, 5, and 6. In these cases, Cols. 36-37 of the Problem Card *must* be blank.

B. Rotation Options

1. Rotation is optional and can be suppressed
2. If rotation is not suppressed, the following series of conditions and options determines the number of variables that will be rotated. (These conditions also determine the number of factors included in the factor matrix.)

The following conditions always hold

  a. A maximum of *19 factors* can be rotated
  b. Only factors with *positive eigenvalues* can be rotated

continued

c. A maximum of 100% of the communality of the original matrix can be accounted for by factors. (In practice, this condition means that, if the first k largest factors account for 100% of the original variance, then *no addition* factors will be calculated.)

The largest number of factors that can be rotated is the minimum of the three values determined by the above conditions. The number of factors to be rotated can be *further* restricted by the following options.

d. Specification of a *maximum number of factors*

e. Specification of the *minimum eigenvalue* to be included among rotated factors. (This specification can be a constant, or can be evaluated as the absolute value of the largest negative eigenvalue)

f. Specification of a *minimum/maximum-factor loading* for factors to be included in the rotation. (In practice, this option means that it is possible to exclude from rotation any factors that have no loadings at or above the specified value. This option should be used with caution because the program will stop calculating factors at the *first* factor that fails this maximum-factor-loading criterion. It sometimes happens that a factor will fail this min/max test when a factor with a smaller eigenvalue would not fail it.)

If any of the communality options d, e, or f are used, the minimum value, so determined, serves as the limit on the number factors to be rotated, if that value is less than the value determined by condition a, b, and c. The options d, e, and f can be suppressed by leaving the appropriate Control-Card columns blank.

REFERENCES

1. Greenstadt, J., "The Determination of the Characteristic Roots of a Matrix by the Jacobi Method," in *Mathematical Methods for Digital Computers*, Ralston and Wilf, Eds. (John Wiley & Sons, Inc., New York, 1959), Chap. 7.

2. Kaiser, H.F., "The Varimax Criterion for Analytical Rotation in Factor Analysis," Psychometrika 23, No. 3 (Sept. 1958).

DESCRIPTIVE TITLE	Program for Analysis of Linear Systems
CALLING NAME	PALS
INSTALLATION NAME	Vogelback Computing Center, Northwestern University
AUTHOR(S) AND AFFILIATION(S)	Dr. James Van Ness Electrical Engineering Department, Northwestern University
LANGUAGE	CDC FORTRAN IV
COMPUTER	CDC 6400
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Lorraine Borman, EIN Technical Representative, Vogelback Computing Center, Northwestern University, 2129 Sheridan Road, Evanston, Ill. 60201 Tel.: (312) 492-3682

FUNCTIONAL ABSTRACT

PALS is actually a series of subroutines combined into one large main program for (1)forming the differential equations of a linear system in matrix form, (2)finding the eigenvalues of the coefficient matrix of the differential equations. and then (3)either computing the loci of the eigenvalues as individual or groups of parameters are varied, finding the eigenvectors of the coefficient matrix, or finding the sensitivities of the eigenvalues to the parameters of the system.

Although PALS was developed initially to study the load-frequency control of a large power system, it is presently written in a general form to make it applicable to many types of systems. The system to be studied may be described in terms of a block diagram or by sets of equations. The method of forming the coefficient matrix of the differential equations from this general input is described in Ref. 1. The only restriction is that the system must be representable by a set of linear differential equations with constant coefficients. Many nonlinear systems, such as the power system, can be studied for small deviations about an operating point by linearizing the system of equations about that operating point.

continued

The accuracy of PALS is dependent on the accuracy of the method used to find the eigenvalues and the eigenvectors of the coefficient matrix of the differential equations. Fortunately, the methods that have been chosen have proven to be very satisfactory. The QR transform developed by Francis[2] is used to find the eigenvalues and is described in Ref. 1. The eigenvectors are found by an inverse-iteration method based on a method described by Wilkinson.[3] After finding the eigenvectors, the program corrects the value found for the eigenvalues by using the Rayleigh quotient. The change in the eigenvalues is insignificant in most cases, but in the few where the accuracy of the QR transform is poor[2] the error is corrected at this step. The inverse-iteration method and the Rayleigh quotient also are described in Ref. 1.

PALS provides two methods of studying the effect of parameter changes on the system eigenvalues. The first is essentially the well-known root-loci approach. Any of the parameters in the system may be incremented through a series of steps, and the resulting change in the eigenvalues will be plotted on a complex plane. The second method finds the sensitivities of specified eigenvalues to the parameter (the derivatives of the eigenvalues to individual or groups of parameters). The sensitivity method thus gives the slope of the eigenvalue locus at a given point. It has the advantage of enabling the user to determine quickly which parameters are important in a given situation. However, the complete locus of the eigenvalue is needed to determine the effect of the parameter over any range of variation. The eigenvalue locus is found by actually incrementing the parameters, whereas the sensitivity is found by using the method described in Ref. 1.

Two versions of PALS are currently being maintained on the CDC 6400 computer. PALS1 is dimensioned for systems up to the 500th order; PALS2 is a simpler version for smaller systems, usually less than 100th order. The actual limits on each of these programs are given in Ref. 1.

REFERENCES

1. Van Ness, J.E., "PALS—A Program for Analyzing Linear Systems," Northwestern Univ. Dept. Elec. Eng. (Mar. 1969). Available from EIN for the cost of duplication and mailing.
2. Francis, J.G.F., "The QR Transformation," Computer J. 4, 265–271 (Oct. 1961); 4, 332–345 (Jan. 1962).
3. Wilkinson, J.H., *The Algebraic Eigenvalue Problem* (Oxford University Press, London, 1965).

000 0045

DESCRIPTIVE TITLE	Statistical Analysis of Single Equation Models
CALLING NAME	QSASE
INSTALLATION NAME	Pennsylvania State University Computation Center
AUTHOR(S) AND AFFILIATION(S)	M.C. Hallberg Department of Agricultural Economics and Rural Sociology Pennsylvania State University
LANGUAGE	FORTRAN IV
COMPUTER	IBM 360/67
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Dr. Daniel L. Bernitt, 105 Computer Building, Pennsylvania State University, University Park, Pa. 16802 Tel.: (814) 865-9527

FUNCTIONAL ABSTRACT

QSASE is a self-contained library program designed to compute ordinary least-squares estimates of single equation regression models such as iterative least-squares estimates of special non-linear regression models (as described in Ref. 1), and two-stage least-squares or limited information maximum likelihood estimates for systems of simultaneous regression equations (as described in Ref. 2). For additional details concerning all options, the user should consult the above references. The following options are available:

1. listing descriptive information about the problem and naming variables
2. using original observations or sums of squares and crossproducts (SSCP) as input data
3. performing various data-editing operations on the input data
4. reading input data from tape or cards
5. printing and/or punching SSCP in either raw form or corrected form
6. printing zero-order correlation matrix

continued

7. printing the mean, standard deviation, standard error of mean, and sum for each variable in the data deck
8. producing a separate listing of all zero-order correlation coefficients that exceed in absolute value a prespecified level
9. producing a scatter diagram of any variable against any other variable
10. stepwise least-squares where a least-squares regression is run and that variable with the smallest t ratio is eliminated in an iterative fashion until all t ratios exceed a prespecified level
11. two-step least squares where a least-squares regression is run and all variables are eliminated whose t ratio do not exceed a prespecified level
12. computing and printing the F ratio for testing the hypothesis that a linear combination of a subset of the k-1 β's for a single-equation model are all zero
13. computing and printing the t ratio for testing the hypothesis that a linear combination of a subset of the k-1 β's for a single-squation model is equal to a specified value
14. computing and/or printing for each observation the value of the dependent variable Y_t, the estimate of Y_t from the computed regression $\hat{Y}_t$, the estimated error $Y_t-\hat{Y}_t$, and the percentage error
15. producing a scatter diagram of the errors $Y_t-\hat{Y}_t$ against any other variable in the data deck
16. two-stage least-squares estimation
17. limited-information maximum-likelihood estimation

REFERENCES

1. Hallberg, M.C., "Statistical Analysis of Single Equation Stochastic Models Using the Digital Computer," Penn. State Univ. Dept. Agric. Econ. & Rural Sociol. Rep. 78 (February 1969). Copies of this report can be obtained through the contact person.

2. Hallberg, M.C., "Statistical Analysis of System of Simultaneous Linear Equations Using the Digital Computer," Penn. State Univ. Dept. Agric. Econ. & Rural Sociol. Rep. 69 (September 1967). Copies of this report can be obtained through the contact person.

DESCRIPTIVE TITLE	PL/I—FORMAC Interpreter
CALLING NAME	FORMAC
INSTALLATION NAME	The Pennsylvania State University Computation Center
AUTHOR(S) AND AFFILIATION(S)	R. Tobey J. Baker R. Crews P. Marks K. Victor International Business Machines Corporation
LANGUAGE	PL/I—FORMAC
COMPUTER	IBM 360/67
PROGRAM AVAILABILITY	Source code available on tape
CONTACT	Dr. Daniel Bernitt, EIN Technical Representative, Computation Center, 105 Computer Building, The Pennsylvania State University, University Park, Pa. 16802 Tel.: (814) 865-9527

FUNCTIONAL ABSTRACT

FORMAC is a system for carrying out formal manipulations on mathematical expressions. This allows for the use of analytic as well as numeric techniques. The most important capability of FORMAC is its accommodation of mathematical expressions as symbolic entities at execution time. For example, the execution of the FORMAC program segment

```
LET( A = X + Y ** 2;
     B = 2;
     C = A/B + 2.8);
```

may be interpreted as assigning the alphanumeric value $(X+Y^2)/2 + 2.8$ to the FORMAC variable C.

FORMAC enables the user to analyze expressions by identifying coefficients, common denominators, lead operators, and some characteristics of the operands. Constants can be factored, left in a rational form, or converted to real notation. New expressions can be synthesized by the simplification, expansion and substitution of

continued

terms, as well as finding the derivatives of functions. The user can specify functions completely or partially in addition to making use of the PL/I functions. Procedures are available for transfering arguments between PL/I and FORMAC program segments.

REFERENCES

Tobey, R., Baker, J., Crews, R., Marks, P., Victor, K., "PL/I-FORMAC Interpreter User's Reference Manual," IBM Publ. 360D 03.3.004 (1967). Copies of the User's Manual will be available through The Pennsylvania State University Computation Center to any user with a PSU account number.

DESCRIPTIVE TITLE	FORMAC Utility Program
CALLING NAME	FMACUT
INSTALLATION NAME	The Pennsylvania State University Computation Center
AUTHOR(S) AND AFFILIATION(S)	Robert Duquet Meteorology Department The Pennsylvania State University H.D. Knoble (assistance) The Pennsylvania State University Computation Center
LANGUAGE	PL/1—FORMAC
COMPUTER	IBM System 360/67
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Dr. Daniel L. Bernitt, EIN Technical Representative, Computation Center, 105 Computer Building, The Pennsylvania State University, University Park, Pa. 16802 Tel.: (814) 865-9527

FUNCTIONAL ABSTRACT

The FORMAC version available at The Pennsylvania State University is an interpretative system; i.e., FORMAC expressions are treated as character strings that are evaluated and executed at run time only. These character strings may be specified by reference to a PL/1 character-string variable. The PL/1 variable, in turn, may be constructed by execution of the PL/1 program in which the FORMAC statements are embedded. FMACUT merely capitalizes on the last-named feature. FORMAC statements are read by the preprocessed, precompiled, and prelink-edited program and are passed to the FORMAC package as character-string arguments. Note that this is made possible also because FORMAC variables need not (in fact *cannot*) be declared in the PL/1 program. Reference 1 should be consulted concerning more-detailed PL/1 FORMAC information.

FMACUT will execute FORMAC statements that are supplied as *input data*. It effectively divorces FORMAC from PL/1 insofar as the user is concerned. This has two advantages: (1) the user need have no knowledge of PL/1 to use FORMAC and (2) FORMAC is available without preprocessing, compilation, or link-editing (which

continued

saves approximately 50 seconds of run time for a null job on the IBM 360/67). The corresponding disadvantage is that PL/1 facilities and those (few) FORMAC statements involving a PL/1 direct interface are not available to the user.

REFERENCES

Tobey, R., Baker, J., Crews, R., Marks, P., and Victor, K., "PL/1 FORMAC Interpreter," IBM Publ. 360D 03.3.004 (1967).

DESCRIPTIVE TITLE	BEEF Mathematical Subroutine
CALLING NAME	BEEFM
INSTALLATION NAME	University of Notre Dame Computation Center
AUTHOR(S) AND AFFILIATION(S)	UNIVAC Division of Sperry Rand Corporation Westinghouse Electric Corporation Baltimore Defense and Space Center Boeing Corporation
LANGUAGE	FORTRAN IV and SLEUTH II
COMPUTER	UNIVAC 1107
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Mrs. Elizabeth Hutcheson, Computer Center, University of Notre Dame, Notre Dame, Ind. 46556 Tel.: (219) 283-7784

FUNCTIONAL ABSTRACT

The BEEF mathematical library is a set of subroutines supplied by UNIVAC to "enhance FORTRAN's abilities as a scientific processor." This enrichment is in the form of subroutines for the evaluation of mathematical functions, matrix arithmetic, and other standard engineering requirements.

See the User Instructions for further information on BEEFM.

REFERENCE

UNIVAC 1107 BEEF Math Routines Manual (UP-3984), UNIVAC Division of Sperry Rand Corporation, (New York, N.Y., March 1965). Available through the UNIVAC 1107 Librarian, Systems Prog. Library Services, UNIVAC Div. of Sperry Rand, Sperry Rand Bldg., New York, N.Y. 10019.

DESCRIPTIVE TITLE	Matching Factor Solutions
CALLING NAME	MATCHFS
INSTALLATION NAME	Office of Data Analysis Research Educational Testing Service (ETS)
AUTHOR(S) AND AFFILIATION(S)	Program due to: R. Pennell, ETS Adaptation due to: J. Barone, ETS
LANGUAGE	F4STAT (Statistically augmented FORTRAN IV)
COMPUTER	IBM 360/65
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Mr. Ernest Anastasio, Office of Data Analysis Research, Educational Testing Service, Rosedale Road, Princeton, N.J. 08540 Tel.: (609) 921-9000 ext. 2552

FUNCTIONAL ABSTRACT

It frequently happens that two factor solutions are obtained in a study, and the question arises as to the extent of similarity. Cliff's[1] approach to the problem is succinctly stated in the abstract of his paper,

> Two problems are considered. The first is that of rotating two factor solutions orthogonally to a position where corresponding factors are as similar as possible. A least-squares solution for transformations of the two factor matrices is developed. The second problem is that of rotating a factor matrix orthogonally to a specified target matrix.

The present program performs least squares, procrustes rotations on a target (T) and a data (D) matrix following the general approach of Cliff.

General Description

The mathematical basis for the program is summarized briefly in the following outline.

continued

Case I: Find an A and B such that transformation (E'E) = minimum

where

$$E = TA - DB \ .$$

A and B are given by

$$T'D = Q$$
$$Q = A\Gamma B'$$

where A , B' and Γ are the Eckart and Young[2] decomposition of Q .

Case II: Find a C such that transformation (E'E) = minimum

where

$$E = T - DC$$

C is given by

$$C = B'A'$$

Case II is the typical target-matching situation were T is held fixed and D rotated to a least-squares fit.

The program can handle matrices as large as 125 rows (variables) X 50 columns.

REFERENCES

1. Cliff, N., "Orthogonal Rotation to Congruence," Psychometrika, *31*, pp. 33–42, (1966).

2. Eckart, C.T., and Young, G., "The Approximation of One Matrix by Another of Lower Rank," Psychometrika, *1*, pp. 211–218, (1936).

3. Pennell, R.J., and Young, F.W., "An IBM System/360 Program for Orthogonal Rotation to Congruence," Behavioral Science, *12*, p. 165, (1967).

000 0061

000 0061

DESCRIPTIVE TITLE	Matrix Decomposition for Points of View Analysis
CALLING NAME	MATDEC
INSTALLATION NAME	Office of Data Analysis Research Educational Testing Service
AUTHOR(S) AND AFFILIATION(S)	J. Ferris Educational Testing Service
LANGUAGE	FORTRAN IV
COMPUTER	IBM 360/65
PROGRAM AVAILABILITY	Deck and listings presently available
CONTACT	Mr. Ernest Anastasio, Off. of Data Anal. Research, Educational Testing Service, Rosedale Road, Princeton, N.J. 08540 Tel.: (609) 921-9000 ext. 2552

FUNCTIONAL ABSTRACT

The purpose of MATDEC is to decompose a rectangular matrix (i.e., a data matrix, X, dimensioned number of variables by number of subjects) into three matrices U, Γ, and W, according to Horst's[1] development (Ref. 1, pp. 364–382), where Γ is a diagonal matrix of eigenvalues and U and W contain the corresponding eigenvectors. It is intended as a first step in Tucker and Messick's approach to an individual differences model for multidimensional scaling.[2]

As is, MATDEC will handle up to 100 variables or subjects, whichever is the lesser dimension of the data matrix. The program uses F4STAT (FORTRAN IV Statistical System) and, in particular, the routine SDGEXT to develop the characteristic roots and vectors of the crossproducts matrix. Further information about F4STAT can be obtained from Mr. Van Hassel, Educational Testing Service, (609) 921-9000, ext. 2557.

REFERENCES

1. Horst, P., *Matrix Algebra for Social Scientists*, (Holt, Rinehart & Winston, New York, 1963).

2. Tucker, L.R., and Messick, S., "An Individual Differences Model for Multidimensional Scaling," Psychometrika, *28*, 333–367 (1963).

000 0061

DESCRIPTIVE TITLE	Factor Analysis by Direct "Oblimin" Method
CALLING NAME	OBLIMIN
INSTALLATION NAME	Educational Testing Service (ETS) Office of Data Analysis Research
AUTHOR(S) AND AFFILIATION(S)	Procedure due to: R. Jennrich, UCLA Program due to: R. Jennrich and H. Harman, ETS Adaptation due to: J. Barone, ETS
LANGUAGE	FORTRAN IV
COMPUTER	IBM 360/65
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Mr. Ernest Anastasio, Office of Data Anal. Research, Educational Testing Service, Rosedale Road, Princeton, N.J. 08540 Tel.: (609) 921-9000 ext. 2552

FUNCTIONAL ABSTRACT

Factor analysis provides that the final solution be in terms of either uncorrelated factors or correlated factors. Beginning in the mid-1940's, following the leadership of Thurstone, there was a trend toward the acceptance of oblique factors. This trend has continued to the present day but, unfortunately, efficient objective means for getting oblique "simple structure" solutions have not generally been available even with modern computers. In 1958, John B. Carroll introduced a whole class of methods for oblique transformation to simple structure. These have come to be known by the term "oblimin" (Ref. 1, pp. 324-326), since they involve *oblique* factors and the *minimization* of a function. The oblimin criterion, which is to be minimized, is given in normalized form by,

$$1. \quad B = \sum_{p<q=1}^{m} \left[n \sum_{j=1}^{n} (v_{jp}^2/h_j^2)(v_{jq}^2/h_j^2) - \gamma \sum_{j=1}^{n} v_{jp}^2/h_j^2 \sum_{j=1}^{n} v_{jq}^2/h_j^2 \right]$$

continued

The v's in these expressions are the elements of the reference-factor structure matrix, i.e., the correlations between the original variables and the reference factors. Actually, what is desired is to have the primary factor pattern exhibit the principles of simple structure, i.e., large values and near-zero values. In the Thurstone school the "reference structure V" is sought that exhibits the simple structure principles, and then the primary factor pattern is obtained by multiplying the matrix V by a diagonal matrix.

The foregoing indirect, and somewhat awkward, procedure has recently been replaced by a direct approach (Ref. 2). Instead of working with the reference factors that are biorthogonal to the primary factors, Jennrich and Sampson set up a criterion for the direct determination of the primary factors that exhibit the simple structure principles. That criterion may be put in the form,

$$2. \qquad F(A) = \sum_{p<q=1}^{m} \left[\sum_{j=1}^{n} a_{jp}^2 a_{jq}^2 - \frac{\delta}{n} \sum_{j=1}^{n} a_{jp}^2 \sum_{j=1}^{n} a_{jq}^2 \right] ,$$

where A is the matrix of primary factor coefficients. Of course, the loadings may be normalized by rows just as in equation 1. An important difference is that the γ in the indirect method ranges between zero and one, while the δ in equation 2 should be zero or negative.

The object of OBLIMIN is to minimize equation 2. The criterion employed is,

$$\frac{F_{i-1} - F_i}{F_0} \leq \varepsilon \quad ,$$

where i is the iteration number. The output is an oblique factor solution satisfying the principles of simple structure, more or less. When δ is equal to zero, the factors are most oblique. For negative values of δ, the factors become less oblique as δ gets smaller. The solution consists of the factor pattern, the correlations among the factors, and the factor structure. At the present time, the program is limited to n = 100 variables and m = 15 factors.

REFERENCES

Harman, H.H., *Modern Factor Analysis*, Second edition, revised, (University of Chicago Press, Chicago, Ill., 1967)

Jennrich, R.I., & Sampson, P.F., "Rotation for Simple Loadings," *Psychometrika*, 1966, *31*, 313-323.

DESCRIPTIVE TITLE	Rao Constellation and Distance Analysis
CALLING NAME	DISCRIM2
INSTALLATION NAME	Michigan State University Computer Institute for Social Science Research
AUTHOR(S) AND AFFILIATION(S)	Stuart Thomas Computer Institute for Social Science Research Michigan State University
LANGUAGE	FORTRAN
COMPUTER	CDC 3600
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Dr. Anders Johanson, Programming Supervisor, Applications Programming, Computer Laboratory, Computer Center, Michigan State University, East Lansing, Mich. 48823 Tel.: (517) 355-4684

FUNCTIONAL ABSTRACT

This program implements the method titled "Constellation and Distance Analysis" when first published by C.R. Rao[1] and later called "Multiple Discriminant Analysis" when presented independently by Bryan[2] and by Lubin. The method operates on a set of variates measured on individuals in several groups. It determines linear combinations of the variates, called discriminant functions, which maximize the ratio of between-group variability to pooled, within-group variability, producing the output listed below. The user can have the program handle data input or write his own subroutine to read data and perform preliminary data manipulations.

Output

1. Job Description
    - User comments
    - Number of groups and variables
    - Options selected

continued

2. Group Summary Statistics
   - Identification
   - Number observations specified on group card
   - Mode of data input and input unit
   - Format statement (if data read by program rather than user)
   - First observation in the group (raw and transformed)
   - Variable means and variances
   - Variable intercorrelation matrix (optional)
3. Overall Data Summary
   - Overall means
   - Overall variances
   - Overall correlation matrix
4. Discriminant Function Information
   - Discriminant criterion
   - Percent trace accounted for by criterion
   - Rao's chi-square and degrees of freedom for the function
   - Function weights for raw data
   - Function weights for data adjusted to unit variances (optional)
   - Mean discriminant score for each group
5. Overall Discriminant Statistics
   - Group centroids in discriminant space (optional)
   - Intercentroid distance matrix (optional)
   - Back solution of discriminant equation (optional)
6. Discriminant Scores (optional)

Capacity

Number of variables must not exceed 30
Number of groups must not exceed 50
Number of observations in any given group must not exceed 9,999
Number of scores (if requested) must not exceed 10 per respondent

REFERENCES

1. Rao, C.R., *Advanced Statistical Methods in Brometric Research*, (John Wiley & Sons, Inc., New York, 1952) Ch. 9.

2. Bryan, J.G., "The Generalized Discriminant Function: Mathematical Foundation and Computational Routine," Harvard Educational Review, *21*, 2, (1951), 90-95.

DESCRIPTIVE TITLE	Nonmetric Multidimensional Scaling
CALLING NAME	KRUSKAL
INSTALLATION NAME	Office of Data Analysis Research Educational Testing Service (ETS)
AUTHOR(S) AND AFFILIATION(S)	Program due to: J.B. Kruskal R.F. Long Bell Telephone Laboratories Modification due to: D. Kirk Educational Testing Service
LANGUAGE	FORTRAN
COMPUTER	IBM 360/65
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Mr. Ernest Anastasio, Office of Data Analysis Research, Educational Testing Service, Rosedale Road, Princeton, N.J. 08540 Tel.: (609) 921-9000 ext. 2552

FUNCTIONAL ABSTRACT

The latest version of the multidimensional scaling program written by J.B. Kruskal of Bell Telephone Labs is available. In addition to the improvements made from earlier versions, a much faster sort has been incorporated into the program and a multi-calculation facility using random starting configurations within one machine run, has been made available. This feature, in addition to being more efficient than the old technique of punching cards for subsequent runs, increases the chances of converging to a global minimum within one machine run with the corresponding savings in time and cost.

A modification to the output was made so that normally only the best of many possible configurations is printed. However, each final configuration and even the individual iterations may be printed (as was done in Kruskal's version) if desired.

The program will handle 60 subjects scaled in up to ten dimensions. Calculations can be repeated up to 99 times on as many as 1800 data values.

continued

REFERENCES

Kruskal, J.B., "Multidimensional Scaling by Optimizing Goodness of Fit to a Nonmetric Hypothesis," Psychometrika, *29*, pp. 1–27, (1964).

Kruskal, J.B., "Nonmetric Multidimensional Scaling: A Numerical Method," Psychometrika, *29*, pp. 115–129, (1964).

000 0071

DESCRIPTIVE TITLE	Simulation Package for University Research and Training
CALLING NAME	SPURT
INSTALLATION NAME	Vogelback Computing Center Northwestern University
AUTHOR(S) AND AFFILIATION(S)	Gustave J. Rath Department of Industrial Engineering and Management Sciences Martin Goldberg Leonard Weiner Northwestern University
LANGUAGE	CDC FORTRAN IV
COMPUTER	CDC 6400
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Lorraine Borman, EIN Technical Representative, Vogelback Computing Center, Northwestern University, 2129 Sheridan Road, Evanston, Ill. 60201 Tel.: (312) 492-3682

FUNCTIONAL ABSTRACT

SPURT is a comprehensive package of USASI Standard FORTRAN routines that are designed for use in simulation modelling. These useful routines, ranging from simple to complex, enable the average FORTRAN programmer to employ simulation techniques without having to learn the semantic and syntactic rules of a new programming language.

The SPURT package is made up of six main parts.

I. *CLOCK Generation—SPURT1*
II. *Stochastic Generators—SPURT2*
III. *Statistical Computations—SPURT3*
IV. *Analog Simulators—SPURT4*
V. *List-Processing and Queue-Manipulation—SPURT5*
VI. *Matrix and Graphical Output—SPURT6*

continued

The following is a listing and brief discussion of the various subroutines contained in each of the six SPURT parts.

CLOCK Generation—SPURT1: to implement discrete-time simulation models; to cause events to occur in the proper time sequence

The CLOCK subroutine consists basically of two lists:

Master Time List—contains events scheduled to happen in the future

Master Time Queue—contains events that could not take place at the time when they were scheduled to and, therefore, have been rescheduled; i.e., they have been blocked and are waiting in a queue.

Events can be stored on either list.

CLOCK recognizes two basic kinds of events:

Exogenous—those that are *external* to the user's routine; these are read from Data Cards by the CLOCK

Endogenous—those that are *internal* to the user's routine; these are generated dynamically and then are maintained by the CLOCK

Stochastic Generators—SPURT2: to generate samples from various probability distributions and to calculate sample values

Subroutine	*Usage*
STOGN1	Permits sampling from a discrete empirical probability distribution defined by the user
STOGN2	Enables the user to approximate a continuous distribution by means of a piecewise linear distribution
UNIFRM	Permits the user to sample real values from a uniform distribution in a defined interval
RANDIN	Provides a uniform distribution of integers in a defined interval
NORMAL	Allows the user to obtain a random sample from a normal distribution with given mean and standard deviation
NEGEXP	Permits the user to obtain a random sample from the negative exponential distribution
POISSN	Provides the user with a random sample from the Poisson distribution

continued

ERLANG	Provides a random sample from the Erlang distribution
DISCRT	Permits sampling from a step function describing a discrete cumulative distribution of integer values
LINEAR	Provides the user with a random sample from a cumulative distribution that is obtained by linear interpolation in a nonequidistant table of real values
DRAW	Provides a boolean value of TRUE or FALSE
RANPER	Generates a uniformly distributed, random permutation of the integers 1, 2, ···, M

Statistical Computations—SPURT3: to calculate statistical parameters and histograms of data arrays

Subroutine	*Usage*
STIX1 STIX2 STIX3	Three interrelated subroutines to accumulate and print out a frequency table and to produce a CalComp plot of a normalized histogram of the table
STIX4	Evaluates the mean, standard deviation, maximum value, and minimum value of an array of real numbers
STIX5	Evaluates the correlation coefficient between two arrays of real numbers
STIX6	Ranks an array of real numbers and produces the median and range of the data within the array
STIX7	Produces a statistical description of the data found in an array, including the sample size, mean, standard deviation, standard error, minimum and maximum values, range, and a printed histogram plot

Analog Simulators—SPURT4: to enable the simulation of analog-computer problems on a digital computer

Subroutine	*Usage*
ANALOG SECND	These two subroutines make it possible to obtain output similar to a hybrid computer

continued

000 0071

List-Processing and Queue-Manipulation—SPURT5: lists are mXn arrays; entries in lists are mX1 arrays.

Subroutine	*Usage*
ADFIFO	Adds an entry at the bottom of the list; it can be removed only after all the elements presently on the list are gone (builds first-in last-out list)
ADLIFO	Adds an entry at the top of the list; it will be removed before any other entry presently on the list (builds last-in first-out list)
REMOVE	Removes the top (or first) entry from a list
PURGE	Destroys the contents of a list
DISPL	Prints the contents of a list

Additional subroutines in SPURT5 provide the capability to rank lists and to delete or to insert entries into lists.

Matrix and Graphical Output—SPURT6: output is facilitated through printing and graphical output

Subroutine	*Usage*
OUT	Prints out a square matrix with column and row headings
NSOUT	Prints out a nonsquare matrix with column and row headings
GRAPH	Produces two-dimensional graphs of plots, using a CalComp plotter

REFERENCES

1. Rath, G.J., "Description of the Simulation Package for University Research and Teaching," Northwestern Univ. Vogelback Computing Ctr., SPURT, Rev. B (1968; unpublished).

DESCRIPTIVE TITLE	Nonmetric Multidimensional Scaling
CALLING NAME	Subroutine KRUSCAL
INSTALLATION NAME	Michigan State University Computer Laboratory
AUTHOR(S) AND AFFILIATION(S)	Alan M. Lesgold Computer Institute for Social Science Research Michigan State University
LANGUAGE	FORTRAN or COMPASS
COMPUTER	CDC 3600
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Dr. Anders Johanson, Programming Supervisor Applications Programming, Computer Laboratory, Computer Center, Michigan State University, East Lansing, Mich. 48823 Tel.: (517) 355-4684

FUNCTIONAL ABSTRACT

KRUSCAL is an implementation of J.B. Kruskal's recently published[1] numerical method for multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis[2]. Given a matrix of similarities or dissimilarities between *n* variables, the routine outputs a configuration of *n* points in a specified number of dimensions such that the distance between any two points is a monotone function of the dissimilarity of the two variables corresponding to those two points.

General Description of Input

KRUSCAL will accept as input any matrix or halfmatrix (below major diagonal) with or without the major diagonal, of similarities or dissimilarities, including, among others, correlation coefficients, confusion probabilities, interaction rates among groups, etc. This matrix need not be symmetric, and the program allows for missing data; interpoint distances corresponding to missing data values do not contribute to the stress.

continued

General Description of Output

The output consists of three sections for the two-dimensional case and two sections for other dimensionalities. In every case, a history of computation is printed, showing for each iteration[1] the following information.

STRESS = Normalized stress achieved[2]

SRAT = The rate of stress improvement = $\frac{\text{STRESS}_i}{\text{STRESS}_{i-1}}$

SRATAV = Weighted average of SRAT

$= \text{SRAT}_i^{1/3} \times \text{SRATAV}_{i-1}^{2/3}$

CAGRGL = Cosine of angle between gradient and previous gradient[1]

COSAV = Weighted average of CAGRGL

$= (1/3)\text{CAGRGL}_i + (2/3)\text{COSAV}_{i-1}$

ACSAV = Weighted average of the magnitude of CAGRGL, computed in the same manner as COSAV

SFGR = Scale factor of gradient, [same as Kruskal's mag(g)][1]

STEP = Step size[1]

Also, in every case, the final number of dimensions is printed. There is an option for printing the interpoint distances in each final configuration, and an option for punching the final configuration onto cards (see description of option cards).

In the two-dimensional case, the final configuration is also plotted (on the line printer, not on the plotter).

REFERENCES

1. Kruskal, J., "Nonmetric Multidimensional Scaling: A Numerical Method", Psychometrika, *29*, pp. 115–129, (June, 1964).

2. Kruskal, J., "Multidimensional Scaling by Optimizing Goodness of Fit to a Nonmetric Hypothesis," Psychometrika, *29*, pp. 1–28, (March, 1964).

DESCRIPTIVE TITLE	FORTRAN Subroutine Package to Solve Ordinary Differential Equations
CALLING NAME	NODE
INSTALLATION NAME	Iowa State University Computation Center
AUTHOR(S) AND AFFILIATION(S)	Iowa State University Computation Center
LANGUAGE	FORTRAN IV
COMPUTER	IBM System 360/65
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Dr. Clair G. Maple, Director, Computation Center, Iowa State University, Ames, Ia. 50010 Tel.: (515) 294-3402

FUNCTIONAL ABSTRACT

NODE is a set of subroutines to solve systems of first-order ordinary differential equations, using discrete-variable methods. The predictor-corrector equations of Crane,[1] which have a wide range of stability, are implemented; the necessary backpoints are initially calculated by using the Runge-Katta-Gill single-step method.[2]

Systems for which NODE is adapted may arise in several ways. Theoretically, every ordinary differential equation of order higher than one can be reduced to a system of first-order equations. For a differential equation of order m, given by

$$y^{(m)} = f[x,y,y',y'',\cdots,y^{(m-1)}],$$

where f is a given function of its m+1 arguments. (Superscripts in parentheses denote derivatives.) The reduction to a first-order system is accomplished by setting

$$y_1 = y,\ y_2 = y',\ \cdots,\ y_m = y^{(m-1)}.$$

If the functions $y_1, y_2, \cdots, y_m$ satisfy the system

$$\begin{aligned} y_1' &= y_2 \\ y_2' &= y_3 \\ &\vdots \\ y_m' &= f(x,y_1,y_2,\cdots,y_m), \end{aligned}$$

continued

then the function $y(x) = y_1(x)$ evidently must satisfy the original equation.

Such first-order systems may also arise in a natural way from many physical problems.

REFERENCES

1. Crane, R.L., "Stability and Local Accuracy of Numerical Methods for Ordinary Differential Equations," Ph.D. thesis, Ia. State Univ. Sci. Technol. (unpublished).

2. Gill,S., "A Process for the Step-by-Step Integration of Differential Equations in an Automatic Digital Computing Machine," Proc. Cambridge Phil. Soc. 47, 96 (1951).

"NODE: A Package of Fortran Subroutines to Solve Ordinary Differential Equations" (user manual), Ia. State Univ. Computation Ctr. (July 1969).

DESCRIPTIVE TITLE	FORTRAN Subroutine Package to Solve Ordinary Differential Equations
CALLING NAME	DNODE
INSTALLATION NAME	Iowa State University Computation Center
AUTHOR(S) AND AFFILIATION(S)	Iowa State University Computation Center
LANGUAGE	FORTRAN IV
COMPUTER	IBM System 360/65
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Dr. Claire G. Maple, Director, Computation Center, Iowa State University, Ames, Ia. 50010 Tel.: (515) 294-3402

FUNCTIONAL ABSTRACT

DNODE is a set of subroutines to solve systems of first-order ordinary differential equations, using discrete-variable methods and double-precision arithmetic. The method of computation is identical to that of the NODE package (EIN No. 000 0082); however, the implementation of double-precision necessitates a slight revision in parameter initialization of the NODE subroutines. (Please refer to the NODE package description.)

REFERENCES

"DNODE: A Package of Fortran Subroutines to Solve Ordinary Differential Equations" (user manual), Ia. State Univ. Computation Ctr. (July 1969).

"NODE: A Package of Fortran Subroutines to Solve Ordinary Differential Equations" (user manual), Ia. State Univ. Computation Ctr. (July 1969). See also the EIN entry for NODE (EIN No. 000 0082).

Crane, R.L., "Stability and Local Accuracy of Numerical Methods for Ordinary Differential Equations," Ph.D. thesis, Ia. State Univ. Sci. Technol. (unpublished).

DESCRIPTIVE TITLE	FORTRAN Subroutine to Solve Simultaneous Linear Equations with Complex Coefficients
CALLING NAME	CGELG
INSTALLATION NAME	Iowa State University Computation Center
AUTHOR(S) AND AFFILIATION(S)	Iowa State University Computation Center
LANGUAGE	FORTRAN IV
COMPUTER	IBM System 360/65
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Dr. Clair G. Maple, Director, Computation Ctr., Iowa State University, Ames, Ia. 50010 Tel.: (515) 294-3402

FUNCTIONAL ABSTRACT

CGELG is a FORTRAN subroutine that solves a set of simultaneous equations with complex coefficients. There exists the capability of handling more than one set of righthand constraints at a time. Yet, the most desirable aspect of the routine is its high degree of accuracy, even on a large, ill-conditioned matrix.

The procedure used is the Gaussian elimination (with complete pivoting) method,[1] modified for complex arithmetic. In this method, the first step is a forward reduction, in which the coefficient matrix is reduced to an upper triangular matrix,using elementary row and column transformations. Then the solution is found by successive back substitution.

REFERENCES

1. Ralston, A., *A First Course in Numerical Analysis* (McGraw—Hill Book Co., Inc., New York, 1965), pp. 339—407.

DESCRIPTIVE TITLE	Minres Method of Factor Analysis
CALLING NAME	MINRES
INSTALLATION NAME	Office of Data Analysis Research Educational Testing Service (ETS)
AUTHOR(S) AND AFFILIATION(S)	Program due to: H. Harman, ETS W. Jones, National Security Agency Y. Fukuda, TRW Systems, Inc. Adaptation due to: J. Barone, ETS
LANGUAGE	FORTRAN IV
COMPUTER	IBM 360/65
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Mr. Ernest Anastasio, Office of Data Analysis Research, Educational Testing Service, Rosedale Road, Princeton, New Jersey 08540 Tel.: (609) 921-9000 ext. 2552

FUNCTIONAL ABSTRACT

The word "minres" is a contraction of "minimum residuals," and designates a method of factor analysis involving the minimization of off-diagonal residuals of a correlation matrix. Such a method, which has long been sought, has many features that recommend it for initial factorization of a correlation matrix. While the objective of the principal-factor method is to extract maximum variance, the objective of the minres method is to "best" reproduce the observed correlations. The latter objective can be traced to Thurstone, "The object of a factor problem is to account for the tests, or their intercorrelations, in terms of a small number of derived variables, the smallest possible number that is consistent with acceptable residual errors" (Ref. 1, p. 61). This problem has been resolved (Ref. 2, Chap. 9) by minimizing the residual correlations (i.e., the differences between the observed values and those reproduced from the factor-analysis model).

The basic factor-analysis model may be put in the form

1. $z_j = a_{j1}F_1 + a_{j2}F_2 + \cdots + a_{jm}F_m + d_jU_j \quad (j = 1,2,\cdots,n)$

continued

in which the common factor loadings in the matrix $A = (a_{jp})$ are the only parameters to be estimated. However such a solution is obtained, the matrix $R^{\dagger}$ of reproduced correlations with communalities in the principal diagonal is given by (Ref. 2, p. 28),

$$2.\quad R^{\dagger} - AA',$$

where uncorrelated factors are assumed without loss of generality. The condition for a least-squares best fit to the off-diagonal correlations may be expressed as follows.

$$3.\quad f(A) = \sum_{k=j+1}^{n} \Sigma n_{j=1}^{n-1} \left[r_{jk} - \sum_{p=1}^{m} a_{jp} a_{kp} \right]^2 = \min \quad .$$

The objective function in (3) is to be minimized under the constraints.

$$4.\quad h_j^2 = \sum_{p=1}^{m} a_{jp}^2 \leq 1 \qquad (j=1,2,\cdots,n) \ .$$

Thus, the object of the minres method is to minimize the function f(A) for a specified number of factors m by varying the values of the factor loadings. The communalities (restricted to numbers between 0 and 1) are obtained as a byproduct of the method.

The basic mathematical method employed in the computer program involves the Gauss-Seidel process whereby successive displacements are introduced in only one row of A at a time, making the objective function f(A) quadratic. The mathematical technique for minimizing a quadratic-functions subject to side conditions involving inequalities is extremely difficult. However, the inequalities can be removed from the problem at hand, so that it can be made tractable (Ref. 2, p. 192).

The computation of factor loadings is continued until the following criterion is met,

$$5.\quad \max_{j,p} \left| {}_{(i)}a_{jp} - {}_{(i-1)}a_{jp} \right| < \varepsilon \qquad (j = 1,\cdots,n;\ \ p = 1,\cdots,m) \ ,$$

when i is the iteration number and ε is preset; or until some maximum number of iterations is reached.

The test statistic, U_m (Ref. 2, p. 197), for testing the significance of m factors is calculated at the conclusion of the program. This is asymptotically dustributed as χ^2 with $\frac{1}{2}[(n - m)^2 + n - m]$ degrees of freedom, where n is the number of variables.

continued

The computer output includes: (1) input parameters and correlation matrix; (2) initial factor matrix (m principal components; alternative is a principal-factor solution dependent on some communality input); (3) minres solution, along with derived communalities; (4) value of the objective function f(A) for this solution; (5) matrix of residual correlations; (6) frequency distribution of the residuals; (7) frequency distribution of the differences in factor loadings between the final iteration and the preceding one; (8) test of significance for the number of common factors; and (9) time for all the preceding calculations.

REFERENCES

Thurstone, L.L., *Multiple-Factor Analysis*, (University of Chicago Press, Chicago, Ill., 1947).

Harman, H.H., *Modern Factor Analysis*, Second Edition, Revised, (University of Chicago, Ill., 1967).

DESCRIPTIVE TITLE	Matrix Operations
CALLING NAME	AES 106
INSTALLATION NAME	The Michigan State University Computer Laboratory, Applications Programming
AUTHOR(S) AND AFFILIATION(S)	William L. Ruble Richard J. Martz Agricultural Experimental Station Michigan State University
LANGUAGE	3600 COMPASS (closed subroutines callable from FORTRAN 600 or COMPASS)
COMPUTER	CDC 3600
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Mr. Anders Johanson, 324 Computer Center, Michigan State University, East Lansing, Mich. 48823 Tel.: (517) 355-4684

FUNCTIONAL ABSTRACT

This is a package of several subroutines for double-precision matrix manipulation.

REFERENCES

Michigan State University Agricultural Experimental Station, *AES MISC Series Description Numbers 106C, 106D, 106E, 106F, 106I, 106J, 106K, 106L.* Mimeographed descriptions available through Mr. Johanson, above, or from EIN at the cost of reproduction.

DESCRIPTIVE TITLE	UCSB On-Line System
CALLING NAME	OLS
INSTALLATION NAME	University of California Santa Barbara UCSB Computer Center
AUTHOR(S) AND AFFILIATION(S)	Dr. David Harris, Principal Investigator Ron Stoughton, Principal Programmer UCSB Computer Research Lab.
LANGUAGE	Basic Assembler Language
COMPUTER	IBM System 360/75
PROGRAM AVAILABILITY	Listing of programs may be obtained by special arrangement with the Computer Research Lab.
CONTACT	Glenn Davis, OLS Manager, Computer Center, Univ. of Cal., Santa Barbara, Cal. 93106 Tel.: (805) 961-2462

FUNCTIONAL ABSTRACT

The UCSB On-Line System (OLS) provides the capability for sophisticated mathematical analysis for use in solving problems where human interaction is either necessary or desired.

OLS accepts both real and complex numbers (scalars) as operands as well as lists of such numbers (vectors). Operations performed on scalars produce scalar results, which can be numerically displayed; operations on vectors produce vector results (the specified operation being performed on each component), and results of computation can be displayed either numerically or graphically. Operands can be stored and used as required. Operators include sine, cosine, logarithm, and exponentiation; and each is executed with a single button push. Facility is provided for interaction between operands of different types (e.g. vectors and scalars). In addition, a limited set of operations manipulate integers used in subscripting.

Additional features are provided to support OLS's basic mathematical capability. Although OLS normally executes each

continued

button as it is pushed, a button sequence may be defined, named, and saved for later execution. Convenient means are provided for editing such sequences. Lists of buttons to be executed can include programmed pauses, allowing manual and programmed activity to be interfaced; as well as branching based upon results of computation. Messages can be composed of alphameric, Greek, and special characters, and displayed. Those characters not specifically provided by OLS may be designed by the user and stored, and then are available for use. A collection of button lists and user-created characters is referred to as a "system". Systems are named and can be permanently stored and later retrieved. Portions of systems may be transferred between systems, and systems may be transferred between users. Sets of scalars and vectors may also be named, permanently stored, and later retrieved.

Apart from OLS's mathematical capability, a recent development provides the ability to create and edit a "deck" of cards and submit it for execution in an OS partition. Operations on string, record and file levels are provided. Data-sets residing on any disk pack within the installation may be fetched, examined, modified, and submitted for execution. Work continues in the general field and further developments are expected.

DESCRIPTIVE TITLE	Smallest Space Analysis	
CALLING NAME	SSA-1	
INSTALLATION NAME	Office of Data Analysis Research Educational Testing Service	
AUTHOR(S) AND AFFILIATION(S)	Procedure due to:	L. Guttman J. Lingoes University of Michigan
	Program due to:	J. Lingoes University of Michigan
	Adaptation due to:	D. Kirk Educational Testing Service
LANGUAGE	MAP	
COMPUTER	IBM 360/65	
PROGRAM AVAILABILITY	Decks and listings presently available	
CONTACT	Mr. Ernest Anastasio, Off. of Data Anal. Research, Educational Testing Service Rosedale Road, Princeton, N.J. 08540 Tel.: (609) 921-9000 ext. 2552	

FUNCTIONAL ABSTRACT

SSA-1 represents a nonmetric technique for finding the smallest Euclidean space for a configuration of points. To quote from the author's description,

> Briefly stated, the problem posed for the...program is: given a matrix of inequalities among pairs of points in a metric or nonmetric space, determine a set of Euclidean coordinates such that the distances calculated from then are a monotonic function of the ranks or order among the inequalities.

(The voluminous list of references related to this topic are available through the contact person.)

According to the authors, a major computational advantage of this program is that it avoids the local minima problem inherent in the Kruskal program (see EIN No. 000 0068) by employing a rank-image principle. The program can handle up to 70 variables in ten dimensions.

DESCRIPTIVE TITLE	Best Least Squares Polynomial Approximation Subroutine
CALLING NAME	LSPOL
INSTALLATION NAME	The University of Iowa University Computer Center
AUTHOR(S) AND AFFILIATION(S)	James Hurt Department of Mathematics University of Iowa
LANGUAGE	FORTRAN IV
COMPUTER	IBM 360/65
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Mrs. Louise R. Levine, Program Librarian, University Computer Center, The Univ. of Iowa, Iowa City, Iowa 52240 Tel.: (319) 353-5580

FUNCTIONAL ABSTRACT

The LSPOL routine computes the coefficients of the polynomial of degree n which best approximates (in the least squares sense) a set of data. The degree of the polynomial must be less than or equal to 20. The method used is that of Beckett and Hurt[1].

REFERENCES

1. Beckett, R., and Hurt, J., *Numerical Calculations and Algorithms*, (McGraw-Hill Book Co., Inc., New York, 1967), pp. 264-290.

DESCRIPTIVE TITLE	Invert Ill-Behaved Matrices Using Hotelling and Bodewig's Iterative Technique
CALLING NAME	HRDMIN
INSTALLATION NAME	The University of Iowa University Computer Center
AUTHOR(S) AND AFFILIATION(S)	Robert G. McCord Graduate Student in Comp. Sci. University of Iowa
LANGUAGE	FORTRAN IV
COMPUTER	IBM 360/65
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Mrs. Louise R. Levine, Program Librarian, Applications Programming, University Computer Center, The Univ. of Iowa, Iowa City, Ia. 52240 Tel.: (319) 353-5580

FUNCTIONAL ABSTRACT

This routine attempts to invert matrices, particularly ill-behaved matrices such as a Hilbert matrix, to a designated degree of accuracy. Bodewig and Hotelling's method of matrix inversion with iteration is used. The matrix to be inverted and an approximation of its inverse are passed to the routine, and it returns the refined inverse and the product of the original matrix and its inverse. An MXM work area must be passed, where M is the dimension in the main program of the matrix to be inverted. Some matrices may be so ill-conditioned that either an approximation of the inverse cannot be found or the inverse cannot be refined to the degree of accuracy specified. In either case the method cannot be used.

Method

A first approximation to the actual inverse is refined to a desired accuracy using Bodewig and Hotelling's technique.

Let A^{-1} be the true inverse to the matrix A, and let A_0^{-1} be a first approximation to A^{-1} such that

continued

$$||I - AA_0^{-1}|| < K < 1$$

where K is a constant, I is the identity matrix, and $||x|| = \max_{1 \leq i \leq n} |x_i|$ is the norm of a matrix X,

$$A_0^{-1} = A^{-1} + E_0$$

where E_0 is an error matrix.

Then $AA_0^{-1} = AA^{-1} + AE_0 = I + AE_0$. Set $R_0 = I - AA_0^{-1}$

$$\begin{array}{ll}
A_1^{-1} = A_0^{-1}(I + R_0) & R_1 = I - AA_1^{-1} \\
A_2^{-1} = A_1^{-1}(I + R_1) & R_2 = I - AA_2^{-1} \\
\vdots & \vdots \\
A_p^{-1} = A_{p-1}^{-1}(I + R_{p-1}) & R_p = I - AA_p^{-1}
\end{array}$$

The norm of the matrix $A_n^{-1} - A^{-1}$ can be made arbitrarily small by increasing the number of iterations since

$$||A^{-1} - A_p^{-1}|| \leq \frac{n^2p||A_0|| \; ||P_0||^{2p}}{1-n \; ||R_0||}$$

where n is the order of A,
p is the number of iterations.

By forcing $||R_0||$ to be less than one, the solution can be made to converge to A^{-1}.

Accuracy

A 24X24 Hilbert matrix was inverted to accuracy of 10^{-4}.

REFERENCES

Householder, A.S., *Principles of Numerical Analysis* (New York: McGraw-Hill Book Co., Inc., 1953), pp. 56-59.

DESCRIPTIVE TITLE	General Purpose Simulation System
CALLING NAME	GPSS
INSTALLATION NAME	Washington University Computing Facilities
AUTHOR(S) AND AFFILIATION(S)	IBM Application Program
LANGUAGE	360 Assembly Language
COMPUTER	IBM 360/50
PROGRAM AVAILABILITY	Proprietary; available for use at Washington University. Available for distribution from IBM.
CONTACT	Dr. C.B. Drebes, Mgr., Scientific Data Processing, Computing Facilities, Box 1098, Washington University, St. Louis, Mo. 63130 Tel.: (314) 863-0100 ext. 3141

FUNCTIONAL ABSTRACT

GPSS is a transaction-oriented language designed for conducting evaluations and experiments concerning the behavior of systems, methods and processes. It has a modular structure which permits "transactions" to flow through the system, where their interactions can be observed and modified. A "clock" is maintained by which events are either scheduled to occur or else determined by one of the eight random number generators provided. Information can be obtained regarding sequencing of operations, scheduling and allocation rules, inventories, queuing disciplines, machine failures, etc. In general, various trade-offs between cost and performance can be studied.

REFERENCES

General Purpose Simulation System/360, Introductory User's Manual (IBM Manual GH20-0304-0), (White Plains, N.Y.: IBM Corp. Tech. Pub. Depart., 1969).

General Purpose Simulation System/360, User's Manual (IBM Manual GH20-0326-3), (White Plains, N.Y.: IBM Corp. Tech. Pub. Depart., 1969).

General Purpose Simulation System/360, OS (360A-CS-17X) Operator's Manual (IBM Manual H20-0311), (White Plains, N.Y.: IBM Corp. Tech. Pub. Depart., 1969).

DESCRIPTIVE TITLE	Continuous System Modeling Program
CALLING NAME	CSMP
INSTALLATION NAME	Washington University Computing Facilities
AUTHOR(S) AND AFFILIATION(S)	IBM Application Program
LANGUAGE	CSMP imbedded in FORTRAN
COMPUTER	IBM 360/50
PROGRAM AVAILABILITY	Proprietary; available for use but not for distribution.
CONTACT	Dr. C.B. Drebes, Mgr., Scientific Data Processing, Computing Facilities, Box 1098, Washington University, St. Louis, Mo. 63130 Tel.: (314) 863-0100 ext. 3141

FUNCTIONAL ABSTRACT

CSMP performs a simulation of a continuous system, obtaining solutions to problems expressed in the form of systems of differential equations or analog block diagrams. Typical applications might be a control engineer's study of the effectiveness of various control system designs, or a study of a cardiovascular system model.

CSMP provides a basic set of functional blocks with which the components of a continuous system may be represented, and it accepts application-oriented statements for defining the connections between these functional blocks. It also accepts FORTRAN IV statements which can be used to handle non-linear and time-variant problems. Input and output are facilitated by means of user-oriented control statements.

REFERENCES

System/360 Continuous System Modeling Program (360A-CX-16X) User's Manual (IBM Manual GH20-0367-3), (White Plains, N.Y.: IBM Corp. Tech. Pub. Depart., 1968).

continued

System/360 Continuous System Modeling Program (360A-CX-16X) Operator's Manual (IBM Manual H20-0368), (White Plains, N.Y.: IBM Corp. Tech. Pub. Depart., 1968).

System/360 Continuous System Modeling Program (360A-CX-16X) Application Description (IBM Manual GH20-0240-2), (White Plains, N.Y.: IBM Corp. Tech. Pub. Depart., 1968).

Operations Research

DESCRIPTIVE TITLE	Generalized Analysis of Transaction Flows
CALLING NAME	INDIFF NUCC047
INSTALLATION NAME	Vogelback Computing Center, Northwestern University
AUTHOR(S) AND AFFILIATION(S)	Steven J. Brams, Syracuse University Converted for the CDC 6400 by John E. Michelsen, Northwestern University
LANGUAGE	CDC FORTRAN IV
COMPUTER	CDC 6400
PROGRAM AVAILABILITY	Deck and listing presently available
CONTACT	Lorraine Borman, EIN Technical Representative, Vogelback Computing Center, Northwestern University, 2129 Sheridan Road, Evanston, Ill. 60201. Tel.: (312) 492-3682

FUNCTIONAL ABSTRACT

The model programmed is the Goodman[1,2] generalization of the Savage-Deutsch[3] transaction-flow model, which is applicable to any kind of quantifiable transactions. The Savage-Deutsch model, programmed by Alker[4] was simplified and extended by Goodman to allow for the analysis of matrices in which some diagonal and/or nondiagonal cells contain blank entries. Such entries, which may either be unknown because data are unavailable or unreliable or be known *a priori* to be equal to zero, can be read in with the nonblank entries. The blank entries are excluded, however, from calculations made by the program for the cells with nonblank entries in the data matrix.

The program calculates three sets of indices for all cells that correspond to those having nonblank entries: (1) the deviation

$$D_{ij} = A_{ij} - E_{ij}$$

continued

between the actual volume of transactions from actor i to actor j, A_{ij} and that predicted or "expected," E_{ij} by the model; (2) the relative deviation, or acceptance, of actor i for actor j's transactions, defined as

$$RA_{ij} = D_{ij}/E_{ij},$$

and (3) the difference between the relative acceptance of actor i for actor j's transactions and the relative acceptance of actor j for actor i's transactions, or

$$RADF_{ij} = RA_{ij} - RA_{ji} = (A_{ij}E_{ji} - A_{ji}E_{ij})/E_{ij}E_{ji},$$

which for all i and j for a symmetrical matrix about the main diagonal, except for sign. This last index provides a measure of the imbalance of transaction flows between two actors, after the effects of size have been removed.

To test the hypothesis that the transaction flows are indifferent to all parameters but the total volume of transactions that each actor sends to and receives from all other actors associated with his nonblank entries, the program calculates the goodness-of-fit statistic

$$\chi^2 = \sum_{i=1}^{N} \sum_{j=1}^{N} D_{ij} RA_{ij} = \sum_{i=1}^{N} \sum_{j=1}^{N} (A_{ij} - E_{ij})^2 / E_{ij},$$

where N is the order of the matrix. The large-sample distribution of this statistic is the χ^2 distribution with L-2N+1 degrees of freedom, where L is the number of cells with nonblank entries.[5] If χ^2 is sufficiently small that the deviations of the observed values from the predicted values could have occurred by chance, then we would accept the hypothesis that the transaction-flow model "explains" the data. When the χ^2 statistic has more than 30 degrees of freedom,

$$ND = (2\chi^2)^{\frac{1}{2}} - (2L-4N+1)^{\frac{1}{2}}$$

may be used as a normal deviate with unit variance to test the fit of the model to the data.

The percentage of transactions that would have to be moved from one cell into another to make the model fit the data perfectly is also calculated by the program. If

$$T = \sum_{i=1}^{N} \sum_{j=1}^{N} A_{ij} = \sum_{i=1}^{N} \sum_{j=1}^{N} E_{ij},$$

continued

the percentage

$$P = \sum_{i=1}^{N} \sum_{j=1}^{N} \frac{50|D_{ij}|}{T}$$

is, in effect, a measure of the overall magnitude of the discrepancies between the observed and predicted proportions for all cells with nonblank entries.[5]

Finally, several statistics summarizing the distribution of the RA and RADF indices for the cells with nonblank entries are incorporated in the program. Adjunct subroutines to the main program calculate the mean, median, standard deviation, and Pearsonian coefficient of skewness of the RA and RADF indices. In addition, those pairs of actors whose D, RA, or RADF indices are greater than a predetermined amount can be specified in the output.

REFERENCES

1. Goodman, L. A., "Statistical Methods for the Preliminary Analysis of Transaction Flows," Econometrica *31*, 197–208 (Jan. – Apr. 1963).
2. Goodman, L. A., "A Short Computer Program for the Analysis of Transaction Flows," Behavioral Sci. *9*, 176–186 (Apr. 1964).
3. Savage, I. R., and Deutsch, K. W., "A Statistical Model of the Gross Analysis of Transaction Flows," Econometrica *28*, 551–572 (July 1960).
4. Alker, H. R., Jr., "An IBM Program for the Gross Analysis of Transaction Flows," Behavioral Sci. *7*, 498–499 (Oct. 1962).
5. Goodman, L. A., "On the Statistical Analysis of Mobility Tables," Amer. J. Sociol. *70*, 564–585 (Mar. 1965).

DESCRIPTIVE TITLE	Linear Programming
CALLING NAME	LINPROG NUCC120
INSTALLATION NAME	Vogelback Computing Center, Northwestern University
AUTHOR(S) AND AFFILIATION(S)	School of Business, Northwestern University
LANGUAGE	CDC FORTRAN IV
COMPUTER	CDC 6400
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Lorraine Borman, EIN Technical Representative, Vogelback Computing Center, Northwestern University, 2129 Sheridan Road Evanston, Ill. 60201 Tel.: (312) 492-3682

FUNCTIONAL ABSTRACT

LINPROG will solve any linear-programming problem with a maximum of 125 variables and 75 constraint equations. It utilizes the original (two-phase) simplex method for solving linear-programming problems and uses a *minimizing* objective function. To solve a maximizing problem, the signs on the objective function coefficients must be reversed before usage.

DESCRIPTIVE TITLE	Critical-Path Summary—Probability Statistic Based on Normal Curve
CALLING NAME	PERTC NUCC022
INSTALLATION NAME	Vogelback Computing Center Northwestern University
AUTHOR(S) AND AFFILIATION(S)	Vogelback Computing Center Northwestern University
LANGUAGE	CDC FORTRAN
COMPUTER	CDC 6400
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Lorraine Borman, EIN Technical Representative, Vogelback Computing Center, Northwestern University, 2129 Sheridan Road, Evanston, Ill. 60201 Tel.: (312) 492-3682

FUNCTIONAL ABSTRACT

PERTC produces a detailed scheduling network, final critical-path summary, and probability statistic for events that are normally distributed about some mean.

REFERENCES

Rogers, H.B., "Construction Project Pert Network," Northwestern Univ. Grad. School Business (Jan. 1964; unpublished). A reference copy may be obtained from the EIN office for the cost of duplicating and mailing.

DESCRIPTIVE TITLE	Synagraphic Computer-Mapping Program
CALLING NAME	SYMAP 5.12
INSTALLATION NAME	Computing Center, The Florida State University
AUTHOR(S) AND AFFILIATION(S)	Howard T. Fisher, Northwestern Technological Institute (overall design and mathematical model) Mrs. O.G. Benson, Northwestern University Computing Center (programming)
LANGUAGE	FORTRAN IV
COMPUTER	Originally developed on IBM 709; also on IBM 360/40 and IBM 7094 at Harvard University Available in EIN through The Florida State University CDC 6400
PROGRAM AVAILABILITY	Decks, listings, and documentation presently available from The Florida State University Computing Center
CONTACT	Ray Soller, Librarian, Computing Center, The Florida State University, Tallahassee, Fla. 32306 Tel.: (904) 599-3418

FUNCTIONAL ABSTRACT

The Synagraphic Computer-Mapping Program (SYMAP) produces maps that depict spatially disposed quantitative and qualitative information. Raw data of every kind (physical, social, economic, etc.) may be related, weighted, and aggregated in a graphic format by assigning values to the coordinate locations of data points or data zones. According to the application and desired representation of data, three basic types of mapping procedure may be specified: contour, conformant, or proximal.

CONTOUR—based on the use of *contour* lines, each of which represents a value remaining constant throughout its length. The map consists of closed curves that connect all points having the same numerical

continued

value. The value at each of the different levels of contour (where a single contour level will represent a particular data value) is determined by the program, according to the scale of the map and the range of the data. Between any two contour lines, a continuous variation is assumed. Therefore, the use of contour mapping should be restricted to the representation of continuously varying information, such as topography, rainfall, or population density.

CONFORMANT—based on the *conformance* to the boundaries of a data zone. This type of mapping is best suited to data for which the representation as a continuously varying surface is inappropriate owing to the significance of physical limits or boundaries. Each predefined data zone is assigned one data value and, depending on its numeric class (range), one representative character on the map itself. Local variation within the zone boundaries will not be apparent, but will, on the average, be correct.

PROXIMAL—based on *proximity* to a data point. In appearance, this type of map is similar to the conformant map. However, point information is used here to define the data zones. Each character location on the output map is assigned the value of the *nearest* data point, using nearest-neighbor techniques. Boundaries are then assumed along the lines where these values change. Then the mapping is carried out as in the conformant type.

REFERENCES

Robertson, J. C., "The SYMAP Programme for Computer Mapping," Cartographic 108–113 (Dec. 1967); taken from a report of the Select Committee on the Ordinance Survey of Scotland.

Fisher, M., "The Laboratory for Computer Graphics," Harvard Univ. Grad. School Design Suppl. (Summer 1967).

Shepard, D., "A Two-Dimensional Interpolation Function for Irregularly Spaced Data," Harvard Univ. Grad. School Design Lab. Computer Graphics (Feb. 1968); available from the Computing Center, The Florida State University. Deals with the subject of analyzing irregularly spaced data derived from a continuous surface. A method is developed for reconstructing the surface from the sampled data. This method is the main device used to generate the maps produced by SYMAP.

DESCRIPTIVE TITLE	Transportation Planning Package
CALLING NAME	TRAN/PLAN
INSTALLATION NAME	The Florida State University Computing Center
AUTHOR(S) AND AFFILIATION(S)	Charles E. Akers William T. Gaupin The Florida State University Computing Center
LANGUAGE	Primary Language: FORTRAN IV Secondary Language: COMPASS Assembly Language
COMPUTER	CDC 6400
PROGRAM AVAILABILITY	Programs generally available through the CDC users group. FSU will make them available to EIN members.
CONTACT	William T. Gaupin, Computing Center, The Florida State University, Tallahassee, Fla. 32306 Tel.: (904) 599-4770

FUNCTIONAL ABSTRACT

The Transportation Planning Package incorporates a set of analytical techniques that enable the user to estimate future transportation requirements and evaluate proposed systems. The collection of survey techniques, analysis method, and computer programs used by TRAN/PLAN have evolved over the past two decades through studies supported by State Highway Departments, U.S. Bureau of Public Roads, and the U.S. Department of Housing and Urban Development. The TRAN/PLAN programs perform the following functions, (1) process data accumulated from inventories or surveys of existing transportation facilities, (2) calibrate and analyze several parameters for a regional transportation network, and (3) predict future intraregional transportation demands. The accuracy of future estimates depends heavily upon the choice of a suitable model, effective intermediate analysis-of-program output, and iterative recalibration of model parameters. Reference 6 explains the overall purpose of transportation planning systems in greater detail.

continued

The TRAN/PLAN Package consists of the following subprograms[1].

(a)	RELINK	Relink nodes of network
(b)	NETWORK	Build and edit network
(c)	NETDVR	Network report generator
(d)	MINPATH	Build minimum paths
(e)	SKIM	Skim (summarize) tape of trees
(f)	COLLECT	Trip data collector
(g)	TRPBLDR	Build trip tables
(h)	ADDER	Add selected trip tables
(i)	SPLITER	Split trips by direction
(j)	UPDATE	Update trip volumes
(k)	MERGE	Merge trip tables
(l)	DEMERGE	Demerge trip tables
(m)	COMPRES	Compress or expand trip tables
(n)	SECTORS	Select sector of network
(o)	TRIPTAB	Trip-table summary generator
(p)	TRIPEND	Trip-end summary generator
(q)	GRAVITY	Gravity model
(r)	FRATAR	Fratar expansion
(s)	ASSIGN	Traffic assignment
(t)	CPACITY	Capacity restraint
(u)	MODSPLT	Modal split model
(v)	CAPRPT	Cross-reference historical-report generator
(w)	CPRSRPT	Historical-report generator
(x)	COMPARE	Statistical comparison of trip volumes
(y)	AVERAGE	Historical-record summary-report generator
(z)	TIMESUM	Time-table-report generator
(aa)	TRPFREQ	Trip frequency distribution

Restrictions which apply to all programs,

1. the maximum allowed zone (centroid) number is 650
2. the maximum allowed node number is 3000
3. the maximum number of tables (purposes) is 4
4. the maximum table number is 4
5. a stacked table may contain a maximum of 8 merged tables

REFERENCES

Akers, C.E. and Gaupin, W.T., "User Manual Transportation Programs for a CDC 6400 Computer," (The Florida State University Computing Center, November 1968; unpublished). Copies of this manual are available through the Librarian, The Florida State University Computing Center.

"Calibrating and Testing a Gravity Model for Any Size Urban Area," Bureau of Public Roads, Washington, D.C., October 1965.

continued

Manual Supplements for the Transportation Planning System, Highway and Local Government Department of South Austrailia, Adelaide, Austrailia, January 1966.

Martin, Brian V., "Minimum Path Algorithms for Transportation Planning," *Research Engineer*, December 1963.

"Modal Split," U.S. Department of Commerce, Bureau of Public Roads, Office of Planning, December 1966.

Shofer, R. and Goodyear, F., "Electronic Computer Applications in Urban Transportation Planning," Proceedings of the 22nd National Conference, Association for Computing Machinery Publication P-67, 1967.

Traffic Assignment Manual, Bureau of Public Roads, Washington, D.C., June 1964.

Transportation Planning System for the Control Data 3600 Computer, Data Centers Division of Control Data Corporation, Minneapolis, Minnesota, May 1965.

Fratar, T.J., "Vehicular Trip Distributions by Successvie Approximation," *Traffic Quarterly*, January 1954, pp. 53-65.

DESCRIPTIVE TITLE	Program for Analysis of Linear Systems
CALLING NAME	PALS
INSTALLATION NAME	Vogelback Computing Center, Northwestern University
AUTHOR(S) AND AFFILIATION(S)	Dr. James Van Ness Electrical Engineering Department, Northwestern University
LANGUAGE	CDC FORTRAN IV
COMPUTER	CDC 6400
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Lorraine Borman, EIN Technical Representative, Vogelback Computing Center, Northwestern University, 2129 Sheridan Road, Evanston, Ill. 60201 Tel.: (312) 492-3682

FUNCTIONAL ABSTRACT

PALS is actually a series of subroutines combined into one large main program for (1)forming the differential equations of a linear system in matrix form, (2)finding the eigenvalues of the coefficient matrix of the differential equations, and then (3)either computing the loci of the eigenvalues as individual or groups of parameters are varied, finding the eigenvectors of the coefficient matrix, or finding the sensitivities of the eigenvalues to the parameters of the system.

Although PALS was developed initially to study the load-frequency control of a large power system, it is presently written in a general form to make it applicable to many types of systems. The system to be studied may be described in terms of a block diagram or by sets of equations. The method of forming the coefficient matrix of the differential equations from this general input is described in Ref. 1. The only restriction is that the system must be representable by a set of linear differential equations with constant coefficients. Many nonlinear systems, such as the power system, can be studied for small deviations about an operating point by linearizing the system of equations about that operating point.

continued

The accuracy of PALS is dependent on the accuracy of the method used to find the eigenvalues and the eigenvectors of the coefficient matrix of the differential equations. Fortunately, the methods that have been chosen have proven to be very satisfactory. The QR transform developed by Francis[2] is used to find the eigenvalues and is described in Ref. 1. The eigenvectors are found by an inverse-iteration method based on a method described by Wilkinson.[3] After finding the eigenvectors, the program corrects the value found for the eigenvalues by using the Rayleigh quotient. The change in the eigenvalues is insignificant in most cases, but in the few where the accuracy of the QR transform is poor[2] the error is corrected at this step. The inverse-iteration method and the Rayleigh quotient also are described in Ref. 1.

PALS provides two methods of studying the effect of parameter changes on the system eigenvalues. The first is essentially the well-known root-loci approach. Any of the parameters in the system may be incremented through a series of steps, and the resulting change in the eigenvalues will be plotted on a complex plane. The second method finds the sensitivities of specified eigenvalues to the parameter (the derivatives of the eigenvalues to individual or groups of parameters). The sensitivity method thus gives the slope of the eigenvalue locus at a given point. It has the advantage of enabling the user to determine quickly which parameters are important in a given situation. However, the complete locus of the eigenvalue is needed to determine the effect of the parameter over any range of variation. The eigenvalue locus is found by actually incrementing the parameters, whereas the sensitivity is found by using the method described in Ref. 1.

Two versions of PALS are currently being maintained on the CDC 6400 computer. PALS1 is dimensioned for systems up to the 500th order; PALS2 is a simpler version for smaller systems, usually less than 100th order. The actual limits on each of these programs are given in Ref. 1.

REFERENCES

1. Van Ness, J.E., "PALS—A Program for Analyzing Linear Systems," Northwestern Univ. Dept. Elec. Eng. (Mar. 1969). Available from EIN for the cost of duplication and mailing.
2. Francis, J.G.F., "The QR Transformation," Computer J. *4*, 265–271 (Oct. 1961); *5*, 332–345 (Jan. 1962).
3. Wilkinson, J.H., *The Algebraic Eigenvalue Problem* (Oxford University Press, London, 1965).

DESCRIPTIVE TITLE	Statistical Analysis of Single Equation Models
CALLING NAME	QSASE
INSTALLATION NAME	Pennsylvania State University Computation Center
AUTHOR(S) AND AFFILIATION(S)	M.C. Hallberg Department of Agricultural Economics and Rural Sociology Pennsylvania State University
LANGUAGE	FORTRAN IV
COMPUTER	IBM 360/67
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Dr. Daniel L. Bernitt, 105 Computer Building, Pennsylvania State University, University Park, Pa. 16802 Tel.: (814) 865-9527

FUNCTIONAL ABSTRACT

QSASE is a self-contained library program designed to compute ordinary least-squares estimates of single equation regression models such as iterative least-squares estimates of special non-linear regression models (as described in Ref. 1), and two-stage least-squares or limited information maximum likelihood estimates for systems of simultaneous regression equations (as described in Ref. 2). For additional details concerning all options, the user should consult the above references. The following options are available.

1. listing descriptive information about the problem and naming variables
2. using original observations or sums of squares and crossproducts (SSCP) as input data
3. performing various data-editing operations on the input data
4. reading input data from tape or cards
5. printing and/or punching SSCP in either raw form or corrected form
6. printing zero-order correlation matrix

continued

7. printing the mean, standard deviation, standard error of mean, and sum for each variable in the data deck
8. producing a separate listing of all zero-order correlation coefficients that exceed in absolute value a prespecified level
9. producing a scatter diagram of any variable against any other variable
10. stepwise least-squares where a least-squares regression is run and that variable with the smallest t ratio is eliminated in an iterative fashion until all t ratios exceed a prespecified level
11. two-step least squares where a least-squares regression is run and all variables are eliminated whose t ratio do not exceed a prespecified level
12. computing and printing the F ratio for testing the hypothesis that a linear combination of a subset of the k–1 β's for a single-equation model are all zero
13. computing and printing the t ratio for testing the hypothesis that a linear combination of a subset of the k–1 β's for a single-squation model is equal to a specified value
14. computing and/or printing for each observation the value of the dependent variable Y_t, the estimate of Y_t from the computed regression $\hat{Y}_t$, the estimated error $Y_t-\hat{Y}_t$, and the percentage error
15. producing a scatter diagram of the errors $Y_t-\hat{Y}_t$ against any other variable in the data deck
16. two-stage least-squares estimation
17. limited-information maximum-likelihood estimation

REFERENCES

1. Hallberg, M.C., "Statistical Analysis of Single Equation Stochastic Models Using the Digital Computer," Penn. State Univ. Dept. Agric. Econ. & Rural Sociol. Rep. 78 (February 1969). Copies of this report can be obtained through the contact person.

2. Hallberg, M.C., "Statistical Analysis of System of Simultaneous Linear Equations Using the Digital Computer," Penn. State Univ. Dept. Agric. Econ. & Rural Sociol. Rep. 69 (September 1967). Copies of this report can be obtained through the contact person.

DESCRIPTIVE TITLE	Static Leontief Input—Output Analysis
CALLING NAME	INOUT
INSTALLATION NAME	The Pennsylvania State University Computation Center
AUTHOR(S) AND AFFILIATION(S)	M.C. Hallberg M. Swope Department of Agricultural Economics, The Pennsylvania State University
LANGUAGE	FORTRAN IV
COMPUTER	IBM System 360/67
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Daniel Bernitt, 105 Computer Building, The Pennsylvania State University, University Park, Pa. 16802 Tel.: (814) 865-9527

FUNCTIONAL ABSTRACT

INOUT will solve the static Leontief input—output problem for a model with as many as 175 sectors. From data consisting of a square input matrix, an output vector, and a set of final demand vectors, a great variety of results may be selectively opted, calculated, and printed. Among these are matrices of technical and interdependency coefficients, matrices of interdependency values and net effects, and some other useful vectors.

REFERENCES

Leontief, W., *et al.*, *Studies in the Structure of the American Economy*, Harvard Economic Research Proj. (Oxford University Press, New York, 1953); theoretical and empirical explorations in input—output analysis.

DESCRIPTIVE TITLE	Synagraphic Mapping
CALLING NAME	SYMAP
INSTALLATION NAME	The Pennsylvania State University Computation Center
AUTHOR(S) AND AFFILIATION(S)	Laboratory for Computer Graphics Harvard University Graduate School of Design Adapted by Larry Rich Jeffrey Simon Larry Sinkey Department of Architectural Engineering, The Pennsylvania State University
LANGUAGE	FORTRAN IV
COMPUTER	IBM System 360/67
PROGRAM AVAILABILITY	Decks, listings, and documentation presently available
CONTACT	Dr. Daniel Bernitt, EIN Technical Representative, 105 Computer Building, The Pennsylvania State University, University Park, Pa. 16802 Tel.: (814) 865-9527

FUNCTIONAL ABSTRACT

The Synagraphic Mapping program (SYMAP) produces maps that depict spatially disposed quantitative and qualitative information. Raw data of every kind (physical, social, economic, etc.) may be related, weighted, and aggregated in a graphic format by assigning values to the coordinate locations of data points or data zones. According to the application and desired representation of data, three basic types of mapping procedure may be specified: contour, conformant, or proximal.

CONTOUR
—based on the use of *contour* lines, each of which represents a value remaining constant throughout its length. The map consists of closed curves that connect all points having the same numerical value. The value at each of the different levels of contour

continued

(where a single contour level will represent a particular data value) is determined by the program, according to the scale of the map and the range of the data. Between any two contour lines, a continuous variation is assumed. Therefore, the use of contour mapping should be restricted to the representation of continuously varying information, such as topography, rainfall, or population density.

CONFORMANT
—based on the *conformance* to the boundaries of a data zone. This type of mapping is best suited to data for which the representation as a continuously varying surface is inappropriate owing to the significance of areal limits or boundaries. Each predefined data zone is assigned one data value and, depending on its numeric class (range), one representative character on the map itself. Local variation within the zone boundaries will not be apparent, but will, on the average, be correct.

PROXIMAL
—based on *proximity* to a data point. In appearance, this type of map is similar to the conformant map. However, point information is used here to define the data zones. Each character location on the output map is assigned the value of the *nearest* data point, using nearest-neighbor techniques. Boundaries are then assumed along the lines where these values change. Then, the mapping is carried out as in the conformant type.

While the contour type of map is mose often used and the easiest to produce, the conformant and proximal maps are often more helpful in the "soft" disciplines. Output is in the form of printed pages that, if the total map size exceeds the width of the computer printed page (13 in.), may easily be glued or pasted together to form a continuous map. Also included in the output is a histogram showing frequencies for given data levels, plus several optional features.

REFERENCES

Robertson, J.C., "The SYMAP Programme for Computer Mapping," Cartographic 108–113 (Dec. 1967); taken from a report of the Select Committee on the Ordinance Survey of Scotland.

Fisher, M., "The Laboratory for Computer Graphics," Harvard Univ. Grad. School Design Suppl. (summer 1967).

continued

Shepard, D., "A Two-Dimensional Interpolation Function for Irregularly Spaced Data," Harvard Univ. Grad. School Design Lab. Computer Graphics (Feb. 1968); available from the EIN Office for the cost of duplication and mailing. Deals with the subject of analyzing irregularly spaced data derived from a continuous surface. A method is developed for reconstructing the surface from the sampled data. This method is the main device used to generate the maps produced by SYMAP.

000 0056

000 0056

DESCRIPTIVE TITLE	UNIVAC 1107 Linear Programming Package
CALLING NAME	LP1107
INSTALLATION NAME	The University of Notre Dame Computing Center
AUTHOR(S) AND AFFILIATION(S)	UNIVAC Division of Sperry Rand
LANGUAGE	Sleuth II/FORTRAN IV
COMPUTER	UNIVAC 1107
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Elizabeth Hutcheson, EIN Technical Representative, Computing Center, Univ. of Notre Dame, Notre Dame, Ind. 46556 Tel.: (219) 283-7784

FUNCTIONAL ABSTRACT

LP1107 is a generalized program for the solution of linear programming problems. More specifically, the LP1107 command structure is a mathematical programming control language; used to prepare a control sequence for a specific programming job.

Either the dual or simplex algorithm, or both, can be used. LP1107 incorporates a true programming language with logical capabilities and full macro capabilities. If a group of commands is being used repetitively, the user may incorporate it into his macro command library and issue the macro command instead. Single and double precision arithmetic are included.

REFERENCES

UNIVAC Systems Programming Library Services, *UP3897 1107LP User's Manual (preliminary)*, (UNIVAC, Sperry Rand Building, New York, 1964).

Aronofsky, J.S., "Growing Applications of Linear Programming with a View Towards Understanding Principles on Which Future LP Systems Must be Based," *Communications of the ACM*, (June 1964).

Copies of these references are available through the local UNIVAC representatives or from the EIN Office at the cost of reproduction and mailing.

000 0056

DESCRIPTIVE TITLE	UNIVAC 1107 PERT/COST, PERT/TIME System
CALLING NAME	PERT (library tape)
INSTALLATION NAME	The University of Notre Dame Computing Center
AUTHOR(S) AND AFFILIATION(S)	UNIVAC Division of Sperry Rand Corporation
LANGUAGE	SLEUTH II and COBOL
COMPUTER	UNIVAC 1107
PROGRAM AVAILABILITY	Listings available
CONTACT	Elizabeth Hutcheson, EIN Technical Representative, Computing Center, University of Notre Dame, Notre Dame, Ind. 46556 Tel. (219) 283-7784

FUNCTIONAL ABSTRACT

The UNIVAC 1107 PERT System provides for integrated time and cost planning and control of research and development programs through implementation of the "work package" costing concept of the Department of Defense/National Aeronautics and Space Administration. The system is composed of two major program modules, PERT/COST and PERT/TIME, which generate reports for the project being analyzed. These reports include activity and event reports, a Work package/Activity listing, a Charge number analysis, etc., from PERT/TIME, and reports on the structure of labor and material costs from PERT/COST.

REFERENCES

PERT/COST General Reference Manual, (UP-4013), (UNIVAC Division of Sperry Rand Corp., N.Y.).

UNIVAC 1107/1108 PERT and CPM: Techniques in the Project Management, (UP 3952), (UNIVAC Division of Sperry Rand Corp., N.Y.).

Available through the local UNIVAC representative or through the EIN Office at the cost of reproduction and mailing.

DESCRIPTIVE TITLE	Simulation Package for University Research and Training
CALLING NAME	SPURT
INSTALLATION NAME	Vogelback Computing Center Northwestern University
AUTHOR(S) AND AFFILIATION(S)	Gustave J. Rath Department of Industrial Engineering and Management Sciences Martin Goldberg Leonard Weiner Northwestern University
LANGUAGE	CDC FORTRAN IV
COMPUTER	CDC 6400
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Lorraine Borman, EIN Technical Representative, Vogelback Computing Center, Northwestern University, 2129 Sheridan Road, Evanston, Ill. 60201 Tel.: (312) 492-3682

FUNCTIONAL ABSTRACT

SPURT is a comprehensive package of USASI Standard FORTRAN routines that are designed for use in simulation modelling. These useful routines, ranging from simple to complex, enable the average FORTRAN programmer to employ simulation techniques without having to learn the semantic and syntactic rules of a new programming language.

The SPURT package is made up of six main parts.

I. *CLOCK Generation—SPURT1*
II. *Stochastic Generators—SPURT2*
III. *Statistical Computations—SPURT3*
IV. *Analog Simulators—SPURT4*
V. *List-Processing and Queue-Manipulation—SPURT5*
VI. *Matrix and Graphical Output—SPURT6*

continued

The following is a listing and brief discussion of the various subroutines contained in each of the six SPURT parts.

CLOCK Generation—SPURT1: to implement discrete-time simulation models; to cause events to occur in the proper time sequence

The CLOCK subroutine consists basically of two lists:

Master Time List—contains events scheduled to happen in the future

Master Time Queue—contains events that could not take place at the time when they were scheduled to and, therefore, have been rescheduled; i.e., they have been blocked and are waiting in a queue.

Events can be stored on either list.

CLOCK recognizes two basic kinds of events:

Exogenous—those that are *external* to the user's routine; these are read from Data Cards by the CLOCK

Endogenous—those that are *internal* to the user's routine; these are generated dynamically and then are maintained by the CLOCK

Stochastic Generators—SPURT2: to generate samples from various probability distributions and to calculate sample values

Subroutine	*Usage*
STOGN1	Permits sampling from a discrete empirical probability distribution defined by the user
STOGN2	Enables the user to approximate a continuous distribution by means of a piecewise linear distribution
UNIFRM	Permits the user to sample real values from a uniform distribution in a defined interval
RANDIN	Provides a uniform distribution of integers in a defined interval
NORMAL	Allows the user to obtain a random sample from a normal distribution with given mean and standard deviation
NEGEXP	Permits the user to obtain a random sample from the negative exponential distribution
POISSN	Provides the user with a random sample from the Poisson distribution

continued

ERLANG	Provides a random sample from the Erlang distribution
DISCRT	Permits sampling from a step function describing a discrete cumulative distribution of integer values
LINEAR	Provides the user with a random sample from a cumulative distribution that is obtained by linear interpolation in a nonequidistant table of real values
DRAW	Provides a boolean value of TRUE or FALSE
RANPER	Generates a uniformly distributed, random permutation of the integers 1, 2, ···, M

Statistical Computations—SPURT3: to calculate statistical parameters and histograms of data arrays

Subroutine	*Usage*
STIX1 STIX2 STIX3	Three interrelated subroutines to accumulate and print out a frequency table and to produce a CalComp plot of a normalized histogram of the table
STIX4	Evaluates the mean, standard deviation, maximum value, and minimum value of an array of real numbers
STIX5	Evaluates the correlation coefficient between two arrays of real numbers
STIX6	Ranks an array of real numbers and produces the median and range of the data within the array
STIX7	Produces a statistical description of the data found in an array, including the sample size, mean, standard deviation, standard error, minimum and maximum values, range, and a printed histogram plot

Analog Simulators—SPURT4: to enable the simulation of analog-computer problems on a digital computer

Subroutine	*Usage*
ANALOG SECND	These two subroutines make it possible to obtain output similar to a hybrid computer

continued

List-Processing and Queue-Manipulation—SPURT5: lists are mXn arrays; entries in lists are mX1 arrays.

Subroutine	*Usage*
ADFIFO	Adds an entry at the bottom of the list; it can be removed only after all the elements presently on the list are gone (builds first-in last-out list)
ADLIFO	Adds an entry at the top of the list; it will be removed before any other entry presently on the list (builds last-in first-out list)
REMOVE	Removes the top (or first) entry from a list
PURGE	Destroys the contents of a list
DISPL	Prints the contents of a list

Additional subroutines in SPURT5 provide the capability to rank lists and to delete or to insert entries into lists.

Matrix and Graphical Output—SPURT6: output is facilitated through printing and graphical output

Subroutine	*Usage*
OUT	Prints out a square matrix with column and row headings
NSOUT	Prints out a nonsquare matrix with column and row headings
GRAPH	Produces two-dimensional graphs of plots, using a CalComp plotter

REFERENCES

1. Rath, G.J., "Description of the Simulation Package for University Research and Teaching," Northwestern Univ. Vogelback Computing Ctr., SPURT, Rev. E (1968; unpublished).

DESCRIPTIVE TITLE	An Algorithm for the Optimization of a Quadratic Form Subject to Linear Restraints
CALLING NAME	ZORILLA
INSTALLATION NAME	Iowa State University Computation Center
AUTHOR(S) AND AFFILIATION(S)	D.J. Soults J.J. Zrubeck V.A. Sposito Statistical Laboratory Iowa State University
LANGUAGE	FORTRAN IV
COMPUTER	IBM 360/65
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Dr. William J. Kennedy, Head, Statistical Numerical Analysis and Data Processing Section, Statistical Laboratory, Iowa State University, Ames, Iowa 50010 Tel.: (515) 294-2260

FUNCTIONAL ABSTRACT

ZORILLA will solve quadratic programming problems on the IBM 360 system. The program is composed of a number of subprograms; each is called by a procedure control card. The sequence of control cards defines the solution procedure.

A manual[1] is available to inform the user about the correct formulation for optimizing a quadratic form subject to linear restrictions, and to provide the user with a detailed explanation of how to use the program (order of data deck, key punching format, control cards, etc.).

The program can be used to minimize numerical problems. Keyword commands such as SCAN or MODEL serve as an aid in finding invalid input data or an incorrectly specified model. The program can scale poorly defined problems upon the use of the SCALE agendum card.

continued

ZORILLA will find the values for $x_1, x_2, \ldots, x_n$ which optimize

$$f(x) = p'x - \tfrac{1}{2}\, x'Cx$$

subject to the conditions that

$$Ax \leq b$$

$$x \geq 0$$

where p, C, A contain specified constants. The function f(x) will be referred to as the function which we are optimizing. C must be an n X n symmetric matrix. [Any quadratic form x'Bx may be expressed by a symmetric matrix $x'(\frac{B+B'}{2})x = \tfrac{1}{2}\, x'Cx$ where $C = (B+B')$.] It is also necessary for C to be a positive (or negative) semi-definite matrix when the objective function f(x) is required to be concave (or convex) in the maximization (or minimization) case.

The program requires a simplex tableau input in the form

$$\begin{array}{cc} & \begin{array}{cc} n & 1 \end{array} \\ \begin{array}{c} n \\ m \end{array} & \begin{bmatrix} -C & -p \\ A & b \end{bmatrix} \end{array}$$

n = number of variables
m = number of restrictions.

The identity matrix need not be entered explicitly into the matrix. It is generated by the program.

The simplex method is utilized for transformation. The procedure for choosing the incoming and outgoing vectors is the procedure developed by Van de Panne and Whinston[2]. The method of transforming and updating vectors is the product form of the inverse based on the revised simplex method.

REFERENCES

1. Soults, D.J., Zrubeck, J.J., and Sposito, V.A., *Stat. Num. Anal. and Data Process. Series, No. 9: ZORILLA Reference Manual*, (Statistical Laboratory, Iowa State University, Ames, Iowa, 1969). Available from the Statistical Laboratory at a cost of $0.75.

2. Van de Panne, C., and Whinston, A., "Simplicial Methods for Quadratic Programming," Naval Research Logistics Quarterly, *11*, pp. 273–302, (1964).

DESCRIPTIVE TITLE	A Regression Model Building System
CALLING NAME	MOUFLON
INSTALLATION NAME	Iowa State University Computation Center
AUTHOR(S) AND AFFILIATION(S)	Bonnie L. Hanson Statistical Laboratory Iowa State University
LANGUAGE	FORTRAN IV
COMPUTER	IBM 360/65
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Dr. William J. Kennedy, Head, Statistical Numerical Analysis and Data Processing Section, Statistical Laboratory, Iowa State University, Ames, Iowa 50010 Tel.: (515) 294-2260

FUNCTIONAL ABSTRACT

MOUFLON provides fast and economical use of computational methods for model building in multiple linear regression. The basic model used is

$$y_i = \beta_0 x_{i0} + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_k x_{ik} + e_i$$

$$i = 1,2, \dots ,n$$

where x_{ij}'s are fixed, y_i's are observed values, and e_i's are assumed to be independent $N(0,\sigma^2)$. The values $x_{i0} \equiv 1$, $i = 1, \dots ,n$, are usually used. If we substitute $\beta_0 = \bar{y} - \beta_1\bar{x}_1 - \dots - \beta_k\bar{x}_k$ into the original model, the result is

$$y_i - \bar{y} = \beta_1(x_{i1} - \bar{x}_1) + \dots + \beta_k(x_{ik} - \bar{x}_k) + e_i$$

or

$$y_i' = \beta_1 x_{i1}' + \dots + \beta_k x_{ik}' + e_i ,$$

which is called the reduced model. The usual procedure solves the reduced normal equations associated with the above model and obtains the constant term by means of the relation

continued

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1\bar{x}_1 - \ldots - \hat{\beta}_k\bar{x}_k \quad .$$

It may be helpful to note that the coefficient matrix corresponding to the reduced normal equations is in fact the matrix of corrected sums of squares and cross products. This coefficient matrix will be called X'X in the future. The following notation is used throughout.

k = total number of independent variables available for inclusion in the regression,

p = number of variables currently included in the regression,

RSS_p = residual sum of squares associated with a p-variate regression

RMS = residual mean sqaure,

REGSS = sum of squares due to regression,

REGMS = regression mean square,

$\hat{\beta}_i^2/c_{ii}$ = reduction in REGSS when ith variable is removed from the regression ($\hat{\beta}_i$ is the estimator of β_i and c_{ii} is the ith diagonal element of $(X'X)^{-1}$

Methods

Four methods are available for selecting the optimal regression.

Hocking and Leslie's Use of the C_p Statistic

Hocking and Leslie's procedure[1] is based on earlier work done by C.L. Mallows.[2,3] Mallows suggested that the selection of a 'good' subset of independent variables in a multiple linear regression be based on the standardized total squared error. He developed the C_p statistic as an estimate of this quantity.

$$C_p = \frac{RSS_p}{\hat{\sigma}^2} - (n - 2p)$$

where $\hat{\sigma}^2$ is the residual mean square obtained by fitting the full model and is used as an estimate of σ^2. Mallows showed that models with small bias tend to have C_p's almost equal to p, and he defined these to be 'good' models. Thus, a graph of C_p versus p will show which of the subsets of independent variates are 'good'. However, this method requires the computation of all regressions.

continued

Hocking and Leslie developed an efficient procedure to find the 'best' (in terms of minimum residual sum of squares) subset or subsets of independent variates. Their approach to the problem is concerned with which subset of r variates, $r = k - p$, should be removed from the regression. It is necessary to note that a minimum reduction in the regression sum of squares due to removing a set of r variables implies that the associated p-variate regression has minimum residual sum of squares. Hocking and Leslie state that an equivalent form of the C_p statistic is

$$C_p = \frac{\text{Red}_p}{\hat{\sigma}^2} + (2p - k)$$

where Red_p is the reduction in regression sum of squares due to removing a set of r variables, where $r = k - p$.

The above equations relate to the reduced model

$$y = \beta_1 + \beta_2 x_2 + \ldots + \beta_k x_k + e$$

where β_1 is the constant term. After adjustment for the notation and model used by the program, the C_p statistic is represented as

$$C_p = \frac{\text{Red}_p}{\hat{\sigma}^2} + (2p - k + 1).$$

The first step of Hocking and Leslie's procedure is to compute the k univariate reductions, θ_i (i.e. θ_i = reduction due to removing the ith variate) and to rank them so that $\theta_1 \leq \theta_2 \leq \cdots \leq \theta_k$. At the same time the variables are relabeled according to the order on the θ_i's. Thus, the removal of the first variable (according to the above labeling) leaves the subset of size $k - 1$ with minimum residual sum of squares among the set of possible subsets of size $k - 1$.

Hocking and Leslie's procedure is based on the following property.

> If the reduction in the regression sum of squares due to eliminating any set of variables for which j is the maximum subscript is not greater than θ_{j+1}, then no subset including any variables with subscripts greater than j can result in a smaller reduction.

The sequential procedure for a given $p = 1, \ldots, k - 2$ (which determine $r = 2, \ldots, k - 1$) is as follows.

continued

Step 1 Compute the reduction R due to removing variables 1,2,...,r (according to above labeling). If $R \leq \theta_{r+1}$, procedure terminates and regression consisting of variables r + 1,...,k is taken to be the 'best' p-variate model. If $R > \theta_{r+1}$, go to Step 2.

Step 2 Include variable r + 1 in the set of variates which are candidates for removal. Then compute reductions for all subsets of size r which include variate r + 1. Find the smallest of all reductions calculated so far (the reduction found in Step 1 is included in the set). If the minimum reduction is not larger than θ_{r+2}, the procedure terminates and the p-variate set corresponding to the minimum reduction is taken to be the 'best' regression of size p. If the minimum reduction exceeds θ_{r+2}, go to Step 3.

Step 3 Add variable r + 2 to the set of variates which are candidates for removal. Compute the reductions for all subsets of size r which include variate r + 2. Find the smallest of all reductions calculated so far (the reductions found in Steps 1 and 2 are included in the set). If the minimum reduction is not larger than θ_{r+3}, the procedure terminates and the p-variate set corresponding to the minimum reduction is taken to be the 'best' regression of size p. If the minimum reduction exceeds θ_{r+3}, go to Step 4.

⋮

The procedure continues through as many steps as are necessary to find the subset with minimum residual sum of squares. After the 'best' subset of size p has been found the value of p is decreased by 1 (r increased by 1) and the procedure is executed once again to find the 'best' subset of size p - 1. Once the 'best' subset of size p' (specified by user) is found, the procedure terminates.

Forward Selection

This method[4] assumes the user has a predetermined order for his independent variate. The variates with highest degree of 'importance' for inclusion in the model must be placed first in the ordering and followed by those of lesser 'importance'.

Often, the user will have a basic set of r, $1 \leq r \leq k$, independent variates which are always needed to predict values of the dependent variable y, and a set of k - r variates which might be of value in the prediction equation. He thus wishes

continued

to find out which of the set of k - r variates should be used to predict y. The forward selection procedure begins with the basic set of r variates and proceeds to test each successive variate by means of the following sequential procedure.

Step 1 Test $H_1 : \beta_{r+1} = 0$
Acceptance implies procedure terminates and the basic set of r variates is used to predict y. Rejection implies procedure continues to Step 2 and the basic set of r variates plus variable r + 1 are included in the model.

Step 2 Test $H_2 : \beta_{r+2} = 0$
Acceptance implies procedure terminates and variates $x_1, x_2, \ldots, x_{r+1}$ are used to predict y. Rejection implies procedure continues to Step 3 and variates $x_1, x_2, \ldots, x_{r+1}, x_{r+2}$ are used to predict y.

Step 3 Test $H_3 : \beta_{r+3} = 0$

.
.
.

The basic test of $H_i : \beta_{r+i} = 0$ is to reject if $u > F_\alpha$ where

$$u = \frac{\hat{\beta}_i^2 / c_{ii}}{RMS}$$

F_α is specified by the user.

Sequential Deletion

This procedure assumes as in Forward Selection[4] that the independent variates' order begins with the 'most important' and ends with the 'least important'. Again, the user specifies a basic set r, $1 \le r \le k$, of independent variables needed to predict values of $\overline{y}$. The sequential deletion procedure begins with the full k-variate model and tests each preceding variate by means of the following sequential procedure.

Step 1 Test $H_0 : \beta_k = 0$
Rejection implies procedure terminates and all k independent variables are used to predict y. Acceptance implies procedure continues to Step 2 and the kth variate is removed from the model.

Step 2 Test $H_1 : \beta_{k-1} = 0$
Rejection implies procedure terminates and $x_1, x_2, \ldots, x_{r+1}$ are used to predict y. Acceptance implies procedure continues to Step 3 and the (k-1)st variate is removed from the model.

continued

Step 3 Test $H_2 : \beta_{k-2} = 0$

$\vdots$

Step k - r Test $H_{k-r-1} : \beta_{r+1} = 0$
Rejection implies procedure terminates and $x_1, x_2, \ldots, x_{r+1}$ are used to predict y. Acceptance implies procedure terminates but x_{r+1} is removed from the model and only $x_1, x_2, \ldots, x_r$ are used to predict y. Again the test of $H_{k-i} : \beta_i = 0$ is to reject if $u > F_\alpha$ where

$$u = \frac{\hat{\beta}_i^2 / c_{ii}}{\text{RMS}} \quad .$$

F_α is specified by the user.

Stepwise Regression

Before we look at the stepwise procedure[5] let us consider a general case. First, let X_1 denote the set of variates $x_1, \ldots, x_p$ which are currently in the model and let X_2 denote the set of variates $x_{p+1}, \ldots, x_k$.

Next, define the sample partial correlation coefficient of the dependent variable y and one of the independent variates from the set X_2, (say x_{p+i}), to be the simple correlation coefficient of y^* and x^*_{p+i}. y^* is the set of residuals resulting from the regression of y on $x_1, x_2, \ldots, x_p$ and x^*_{p+i} is the set of residuals resulting from the regression of x_{p+i} on $x_1, x_2, \ldots, x_p$. Let the sample partial correlation coefficient of y and x_{p+i} be denoted by r_{p+i}.

Now consider the case of transferring one of the variates, x_{p+j}, from set X_2 to set X_1 , (i.e. including x_{p+j} in the regression equation). Let RSS_{p+j} denote the residual sum of squares for the regression of y on $x_1, x_2, \ldots, x_p, x_{p+j}$. Note that the value of j can be $1, 2, \ldots, k-p$.

Let x_{p+i} be 'the' variable transferred to the set X_1 where i satisfies

$$RSS_{p+i} \leq RSS_{p+j} \qquad j = 1, 2, \ldots, k-p$$

This is equivalent to

$$r^2_{p+i} \geq r^2_{p+j} \qquad j = 1, 2, \ldots, k-p$$

since it can be shown that

$$RSS_{p+j} = RSS_p (1 - r^2_{p+j}) \quad .$$

continued

Thus, the selection of the largest r^2_{p+j} for $j = 1,2,\ldots,k\text{-}p$, minimizes RSS_{p+j}.

Now consider the case of transferring one of the independent variables x_j from the set X_1 to the set X_2 (i.e. removing it from the regression). Let RSS_j denote the residual sum of squares for the regression of y on $x_1,x_2,\ldots,x_{j-1},x_{j+1},\ldots,x_p$. It can be shown that

$$RSS_j = RSS_p + \frac{\hat{\beta}^2_j}{c_{jj}} \quad .$$

Let x_i be 'the' variable transferred to the set X_2 where i satisfies

$$\frac{\hat{\beta}^2_i}{c_{ii}} \le \frac{\hat{\beta}^2_j}{c_{jj}} \qquad j = 1,2,\ldots,p$$

or

$$(t^2_c)_i \le (t^2_c)_j \qquad j = 1,2,\ldots,p$$

where $(t^2_c)_i = \dfrac{\hat{\beta}^2_i/c_{ii}}{RMS}$. Thus, the selection of the smallest $(t^2{}_c)_j$ for $j = 1,2,\ldots,p$ minimizes RSS_j.

The above considerations give way to the stepwise regression procedure, which consists of two alternating steps and examination of termination criteria after each step. The procedure terminates when any one of the following criteria is encountered.

1. There is no variable to enter and no variable to remove.
2. The procedure dictates that the same variable be entered and removed successively. This can be corrected by changing the F levels if the user so wishes.
3. The total number of steps executed reaches the maximum number of steps specified by the user.

The procedure begins with Step 1 and no variables entered in the model.

continued

Step 1 Enter variable i into the regression if i satisfies

$$r^2_{p+i} \geq r^2_{p+j} \quad j = 1,2,\ldots,k-p \quad \text{and} \quad (t^2_c)_{p+i} = \frac{(n-p-2)r^2_{p+i}}{(1-r^2_{p+i})} > F_{in} ,$$

where F_{in} is the F level to enter a variable and is specified by the user.

The termination criteria are now checked. If any one of the three criteria is satisfied, the program stops computations. If none of the criteria are satisfied, the program continues to Step 2.

Step 2 Remove variable i from the regression if i satisfies

$$(t^2_c)_i \leq (t^2_c)_j \quad j = 1,2,\ldots,p \quad \text{and} \quad (t^2_c)_i < F_{out} ,$$

where F_{out} is the F level to remove a variable and is specified by the user.

The termination criteria are now checked. If any one of the three criteria are satisfied, the program stops. If none of the criteria are satisfied, the program returns to Step 1.

REFERENCES

1. Hocking, R.R., and Leslie, R.N., "Selection of the Best Subset in Regression Analysis," Technometrics, *9*, 1967, pp. 531–540.

2. Mallows, C.L., *Choosing Variables in a Linear Regression: a Graphical Aid.* Paper presented at the Cent. Reg. Meeting of the Inst. of Math. Stat., Manhattan, Kansas, 1964.

3. Mallows, C.L., *Choosing a Subset Regression.* Paper presented at the Joint Stat. Meeting, Los Angeles, Cal., 1966.

4. Larson, H.J., *Sequential Model Building for Prediction in Regression Analysis*, unpublished Ph.D. thesis, (Iowa State University Library, Ames, Iowa, 1960).

5. Hemmerle, W.J., *Statistical Computations on a Digital Computer*, (Blaisdell Publishing Company, a Division of Ginn and Company, Waltham, Mass., 1967).

continued

6. Garside, M.J., "The Best Subset in Multiple Regression Analysis," Appl. Stat. J. of the Roy. Stat. Soc., *14*, Series C, 1965.

7. Hanson, B.L., *Stat. Num. Anal. and Data Process. Series, No. 12: MOUFLON Reference Manual*, (Statistical Laboratory, Iowa State University, Ames, Iowa, 1969). Available from the Statistical Laboratory at a cost of $0.75.

DESCRIPTIVE TITLE	Synagraphic Mapping Program
CALLING NAME	SYMAP
INSTALLATION NAME	Washington University Computing Facilities
AUTHOR(S) AND AFFILIATION(S)	Harvard University Graduate School of Design
LANGUAGE	FORTRAN IV
COMPUTER	IBM 360/50
PROGRAM AVAILABILITY	Proprietary; available for use but not distribution
CONTACT	Dr. Charles Drebes, Manager, Scientific Data Processing, Computing Facilities, Box 1098, Washington University, St. Louis, Mo. 63130 Tel.: (314) 863-0100 ext. 3141

FUNCTIONAL ABSTRACT

SYMAP is a computer program designed to allow city planners, geographers and others to produce low cost graphic displays of spatial patterns using standard computer line printers, by producing maps which graphically depict spatially disposed quantitative and qualitative information. It is suited to a broad range of applications, and is provided with numerous options to meet widely varying requirements. Raw data of every kind (physical, social, economic, etc.) when given to the computer may be related, manipulated, weighted, and aggregated in any manner desired. By assigning values to the coordinate locations of data points or data zones, one or more of three types of map may be produced, as specified by the user: conformant (choropleth), contour, and proximal. Potential applications are independent of the scale at which one wishes to display data. Studies (at other universities) have included a living cell, land parcels, blocks, tracts, towns, states, and continents. In each case, a common factor was the spatial distribution of a variable and a need to display the patterns associated with this distribution.

REFERENCES

Reference Manual for Synagraphic Computer Mapping—"SYMAP" (Cambridge, Mass.: Harvard Univ. Grad. Sch. of Design, Comp. Graph. Lab.).

DESCRIPTIVE TITLE	FORTRAN Subroutine to Solve Simultaneous Linear Equations with Complex Coefficients
CALLING NAME	CGELG
INSTALLATION NAME	Iowa State University Computation Center
AUTHOR(S) AND AFFILIATION(S)	Iowa State University Computation Center
LANGUAGE	FORTRAN IV
COMPUTER	IBM System 360/65
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Dr. Clair G. Maple, Director, Computation Ctr., Iowa State University, Ames, Ia. 50010 Tel.: (515) 294-3402

FUNCTIONAL ABSTRACT

CGELG is a FORTRAN subroutine that solves a set of simultaneous equations with complex coefficients. There exists the capability of handling more than one set of righthand constraints at a time. Yet, the most desirable aspect of the routine is its high degree of accuracy, even on a large, ill-conditioned matrix.

The procedure used is the Gaussian elimination (with complete pivoting) method,[1] modified for complex arithmetic. In this method, the first step is a forward reduction, in which the coefficient matrix is reduced to an upper triangular matrix, using elementary row and column transformations. Then the solution is found by successive back substitution.

REFERENCES

1. Ralston, A., *A First Course in Numerical Analysis* (McGraw—Hill Book Co., Inc., New York, 1965), pp. 339—407.

DESCRIPTIVE TITLE	Mathematical Programming System
CALLING NAME	MPS/360
INSTALLATION NAME	Washington University Computing Facilities
AUTHOR(S) AND AFFILIATION(S)	IBM Application Program
LANGUAGE	360 Assembly Language
COMPUTER	IBM 360/50
PROGRAM AVAILABILITY	Proprietary: available for use at Washington University. Available for distribution from IBM.
CONTACT	Dr. C. B. Drebes, Mgr., Scientific Data Processing, Computing Facilities, Box 1098, Washington University, St. Louis, Mo. 63130 Tel.: (314) 863-0100 ext. 3141

FUNCTIONAL ABSTRACT

Linear programming (LP) is a mathematical technique designed to analyze the potentialities of alternate business activities and to choose those that permit the best use of resources in the pursuit of a desirable objective. It has many uses. For example, it can analyze capital, raw materials, manpower, plant and storage facilities and then translate its findings into minimum costs and maximum profits for its user; it can be used to allocate, assign, schedule, select or evaluate whatever possibilities limited resources possess for different jobs; it can blend, mix, distribute, control, order, budget, bid, cut, trim, price, purchase, and plan; it can deduce the most profitable method of transporting goods from plant to warehouse to outlet.

MPS/360 is composed of a set of procedures, a subset of which deals only with LP. The strategy for solving an LP problem is the ordered execution of a series of these procedures. A second set of procedures handles separable programming problems.

The user conveys the proposed strategy to MPS via the MPS control language. The procedure call statement of the control

continued

language calls the LP procedures and transfers arguments to them. MPS control statements are preprocessed by the control program COMPILER.

A catalogued procedure is available for simple linear programming models, reducing the user's work to merely data arrangement.

REFERENCES

Mathematical Programming System/360, Application Description (IBM Manual GH20-0136-3), (White Plains, N.Y.: IBM Corp. Tech. Pub. Depart., 1969).

Mathematical Programming System/360 (360A-CO-14X), Version 2, Control Language User's Manual (IBM Manual GH20-0290-3), (White Plains, N.Y.: IBM Corp. Tech. Pub. Depart., 1969).

Mathematical Programming System/360 (360A-CO-14X), Version 2, Linear and Separable Programming User's Manual (IBM Manual GH20-0476-0), (White Plains, N.Y.: IBM Corp. Tech. Pub. Depart., 1969).

Mathematical Programming System/360 (360A-CO-14X), Message Manual (IBM Manual GH20-0603-0), (White Plains, N.Y.: IBM Corp. Tech. Pub. Depart., 1969).

Copies of these manuals can be obtained through IBM branch offices.

DESCRIPTIVE TITLE	General Purpose Simulation System
CALLING NAME	GPSS
INSTALLATION NAME	Washington University Computing Facilities
AUTHOR(S) AND AFFILIATION(S)	IBM Application Program
LANGUAGE	360 Assembly Language
COMPUTER	IBM 360/50
PROGRAM AVAILABILITY	Proprietary; available for use at Washington University. Available for distribution from IBM.
CONTACT	Dr. C.B. Drebes, Mgr., Scientific Data Processing, Computing Facilities, Box 1098, Washington University, St. Louis, Mo. 63130 Tel.: (314) 863-0100 ext. 3141

FUNCTIONAL ABSTRACT

GPSS is a transaction-oriented language designed for conducting evaluations and experiments concerning the behavior of systems, methods and processes. It has a modular structure which permits "transactions" to flow through the system, where their interactions can be observed and modified. A "clock" is maintained by which events are either scheduled to occur or else determined by one of the eight random number generators provided. Information can be obtained regarding sequencing of operations, scheduling and allocation rules, inventories, queuing disciplines, machine failures, etc. In general, various trade-offs between cost and performance can be studied.

REFERENCES

General Purpose Simulation System/360, Introductory User's Manual (IBM Manual GH20-0304-0), (White Plains, N.Y.: IBM Corp. Tech. Pub. Depart., 1969).

General Purpose Simulation System/360, User's Manual (IBM Manual GH20-0326-3), (White Plains, N.Y.: IBM Corp. Tech. Pub. Depart., 1969).

General Purpose Simulation System/360, OS (360A-CS-17X) Operator's Manual (IBM Manual H20-0311), (White Plains, N.Y.: IBM Corp. Tech. Pub. Depart., 1969).

DESCRIPTIVE TITLE	Continuous System Modeling Program
CALLING NAME	CSMP
INSTALLATION NAME	Washington University Computing Facilities
AUTHOR(S) AND AFFILIATION(S)	IBM Application Program
LANGUAGE	CSMP imbedded in FORTRAN
COMPUTER	IBM 360/50
PROGRAM AVAILABILITY	Proprietary; available for use but not for distribution.
CONTACT	Dr. C.B. Drebes, Mgr., Scientific Data Processing, Computing Facilities, Box 1098, Washington University, St. Louis, Mo. 63130 Tel.: (314) 863-0100 ext. 3141

FUNCTIONAL ABSTRACT

CSMP performs a simulation of a continuous system, obtaining solutions to problems expressed in the form of systems of differential equations or analog block diagrams. Typical applications might be a control engineer's study of the effectiveness of various control system designs, or a study of a cardiovascular system model.

CSMP provides a basic set of functional blocks with which the components of a continuous system may be represented, and it accepts application-oriented statements for defining the connections between these functional blocks. It also accepts FORTRAN IV statements which can be used to handle non-linear and time-variant problems. Input and output are facilitated by means of user-oriented control statements.

REFERENCES

System/360 Continuous System Modeling Program (360A-CX-16X) User's Manual (IBM Manual GH20-0367-3), (White Plains, N.Y.: IBM Corp. Tech. Pub. Depart., 1968).

continued

System/360 Continuous System Modeling Program (360A-CX-16X) Operator's Manual (IBM Manual H20-0368), (White Plains, N.Y.: IBM Corp. Tech. Pub. Depart., 1968).

System/360 Continuous System Modeling Program (360A-CX-16X) Application Description (IBM Manual GH20-0240-2), (White Plains, N.Y.: IBM Corp. Tech. Pub. Depart., 1968).

DESCRIPTIVE TITLE	A Primal-Dual Transportation Algorithm
CALLING NAME	TRANSPRT
INSTALLATION NAME	Washington University Computing Facilities
AUTHOR(S) AND AFFILIATION(S)	D.E. Burlingame Washington University Computing Facilities
LANGUAGE	FORTRAN IV
COMPUTER	IBM 360/50
PROGRAM AVAILABILITY	Deck and listing presently available
CONTACT	Dr. C.B. Drebes, Mgr., Scientific Data Processing, Computing Facilities, Box 1098, Washington University, St. Louis, Mo. 63130 Tel.: (314) 863-0100 ext. 3141

FUNCTIONAL ABSTRACT

TRANSPRT solves a standard transportation problem using the primal-dual transportation algorithm. This program solves a model in which the objective is to "transport" a single commodity from various origins to different destinations at a minimum total shipping cost. The availability at each origin, the demand at each destination, and the cost to ship one unit of the project from any origin to any destination are required inputs to this model. This model can be applied to certain other types of industrial or business problems that have nothing to do with shipping. Personnel assignment, machine assignment, product and inventory scheduling are a few such applications.

The program utilizes integer arithmetic and requires integer input. If the sum of the origin availabilities is not equal to the sum of the destination requirements, an artificial origin or destination is set up, with zero costs, to handle the excess. The program will handle up to 50 origins and 150 destinations, using 94K bytes of core storage.

REFERENCES

Hadley, G., *Linear Programming* (Reading, Mass: Addison-Wesley Pub. Co., Inc., 1962).

Physical Sciences

DESCRIPTIVE TITLE	Structure-Factor Least-Squares Refinement Program for IBM 7090
CALLING NAME	ORFLS-PX
INSTALLATION NAME	University of Pittsburgh The Computer Center
AUTHOR(S) AND AFFILIATION(S)	Dr. Ryonosuke Shiono Crystallography Laboratory University of Pittsburgh
LANGUAGE	FORTRAN II, except SMI, which is in FAP(UMAP).
COMPUTER	IBM 7090 with 32K core storage 2–3 magnetic-tape units in addition to system input/output.
PROGRAM AVAILABILITY	Proprietary; usage permitted, but program deck or listing not available
CONTACT	Dr. Ryonosuke Shiono, Crystallography Laboratory, 300 Thaw Hall, University of Pittsburgh, Pittsburgh, Pa. 15213 Tel.: (412) 621-3500 ext. 7124

FUNCTIONAL ABSTRACT

ORFLS-PX is an extensively modified version of the FORTRAN least-squares program by Busing, Martin, and Levy. Their program,

> performs successive cycles of refinement using the full matrix of the normal equations. The parameters which may be adjusted include several scale factors, an over-all temperature factor coefficient, the neutron scattering factors, individual atom multipliers, atomic coorinates, and isotropic or anisotropic individual atom temperature factor coefficients. The parameters to be varied may be specified arbitrarily, and structures of any symmetry may be accommodated. The refinement may be based either on the structure factors or their squares, and the observations may be weighted individually, or the use of unit weights may be specified.[1]

Features of ORFLS-PX

1. Reflexion-data input may be from cards or from magnetic tape

continued

2. Uses atomic-scattering-factor table of f versus $\sin\theta$
3. External scale factors now refer to F_{obs}', rather than to F_{cal}'s
4. Weighting scheme may be changed or weights modified by a subroutine
5. Limitations on maximum number of atoms, parameters to be used; number of parameters to be refined have been enlarged
6. Isotropic and anisotropic temperature factors for atoms may be used in mixed fashion
7. Anomalous dispersion correction may be made[2]
8. Calculated structure factors, together with the scales F_{obs} are written on a binary tape, which may be used as an input for the Fourier program or for a new cycle of the least-squares program.[3]
9. Using an output tape as the input, a part of the structure may be added as fixed contribution for a partial-mode calculation, (this either saves time in calculation or extends the limit of number of atoms to be used)
10. Different set of parameters may be refined in successive cycles of full-matrix refinement, when not all the parameters can be refined at the same time

General Description of the Program Setup

To provide more storage space for the parameter refinement, ORFLS-PX is divided into three separate core-load packages: MAIN(1), MAIN(2), and MAIN(3).

MAIN(1) reads the reflexion data on cards or on tape, interpolates the atomic-scattering factors for each kind, and assigns the weight according to the specification. This part may be bypassed if a tape from a previous cycle of calculation is used as the input. Otherwise, the program produces an intermediate output tape.

MAIN(2) reads back the MAIN(1) tape, one reflexion at a time and calculates structure factor and the necessary derivatives for least-squares refinement. The calculated structure factors and scaled F_{obs}' are on the output binary tape. This tape is rewound at the end of every cycle so that the data on this tape are the result of latest calculation. This tape may be used as the input for fixed contribution in a partial-mode calculation.

MAIN(3) prints out the correlation matrix and other information on various agreements at the end of all calculation.

continued

Card formats used for reflexion data and atomic parameters and for symmetry operations are identical to many other interconnected programs on the IBM 7090 and IBM 1130 as well as the IBM 1620. Usually, the user has only to punch a few Control Cards to run any one of the programs.

The Data and Parameter Cards are identical to those in the Block-Diagonal Least-Squares programs[3] on the IBM 7090 and 1130, and so the user may easily switch from one program to another.

The reflexion-data file on magnetic tape has, again, the same format between this and Block-Diagonal, E, Fourier and other programs on the IBM 7090 (see Ref. 3), and it may be used interchangeably.

ORFLS-PX is an independent program, as compared to some others incorporated in a big system, permitting one to write a subroutine to fit a particular problem, and replace the dummy variables provided in the program by one's own.

REFERENCES

1. Busing, W.R., Martin, K.O., and Levy, H.A., "A FORTRAN Crystallographic Least-Squares Program," ORNL Tech. Mem. No. TM305 (Oak Ridge National Laboratories, Oak Ridge, 1962). The modified version of this program is used for calculations that involve scattering data and crystallographic structures. Additional information can be obtained through the contact person.
2. Ibers, J.A., and Hamilton, W.C., "Dispersion Correction and Crystal Structure Refinement," Acta Cryst. *17*, (1964), 781-782.
3. Additional information can be obtained through the contact person.

DESCRIPTIVE TITLE	Information Retrieval Services
CALLING NAME	
INSTALLATION NAME	University of Georgia Computer Center
AUTHOR(S) AND AFFILIATION(S)	Information Science Group University of Georgia
LANGUAGE	Assembler Language & PL/1
COMPUTER	IBM 360/65
PROGRAM AVAILABILITY	Proprietary; available for use but not for distribution
CONTACT	Miss Margaret Park, Supervisory Inform. Scientist, Computer Center, The Univ. of Ga., Athens, Ga. 30601 Tel.: (404) 542-3741

FUNCTIONAL ABSTRACT

In any endeavor, scientists need to keep constantly abreast of activities in their field of interest, to be on the lookout for new ideas, and to maintain a library of useful references. The proliferation of new scientific knowledge is rapidly outpacing the capabilities of conventional information-handling and publishing techniques. Scientists are now turning to computer-based methods to help speed and channel the flow of information on a timely basis.

The Computer Center at the University of Georgia is actively engaged in establishing an Information Center. Mechanized data bases from several scientific organizations are presently available and in use at the Center. Subject areas currently represented are biology, biochemistry, nuclear science, and chemistry, including structural data files for chemical compounds. Other tape services in fields such as medicine, engineering, physics, geology, etc., will be added as interst is expressed in these subject areas.

The Computer Center's Information Sciences Unit offers assistance in creating search profiles and current awareness and retrospective searches of the scientific literature.

continued

The staff of the Computer Center includes highly trained professionals in practically every subject discipline. In addition to chemists and a microbiologist in the Information Sciences Unit, the Center also has full time staff with specialties in physics, engineering, statistics, biology, and forestry, with consulting staff available in other disciplines. All staff are also familiar with computer systems and their applications.

The Computer Center is constantly seeking better ways to satisfy the information requirements of the scientific community and we will greatly appreciate any suggestions made in this direction. New services will be added as rapidly as possible to meet newly identified needs and uses.

Persons desiring to use the Information Retrieval Services are directed to the contact person.

Statistics & Meas.

DESCRIPTIVE TITLE	Mann-Whitney U Test
CALLING NAME	MANNWH NUCC085
INSTALLATION NAME	Vogelback Computing Center, Northwestern University
AUTHOR(S) AND AFFILIATION(S)	Alan Lupa, Department of Chemical Engineering, Northwestern University Dennis R. Goldenson, Department of Political Science, Northwestern University
LANGUAGE	CDC FORTRAN IV
COMPUTER	CDC 6400
PROGRAM AVAILABILITY	Deck and listing presently available
CONTACT	Lorraine Borman, EIN Technical Representative, Vogelback Computing Center, Northwestern University, 2129 Sheridan Road, Evanston, Ill. 60201. Tel.: (312) 492-3682

FUNCTIONAL ABSTRACT

This program calculates the Mann-Whitney U statistic and reports its significance at one of five levels of confidence (that is, accept H_0, $P > .10$; reject H_0, $P < .10$, $P < .05$, $P < .02$, and $P < .01$). There is no limit to the number of problems that can be handled. Each subsample in a problem can have up to 2,000 cases. This is purely arbitrary; if one would want a larger sample-size capacity, he could merely alter the DIMENSION statements in the source deck to the limits of core storage. Printed output gives the problem number, alphanumeric problem label, the sample sizes, the input variable format, the median and decile range for each group, the U-test statistic (the smaller value), the value of Z where the larger group is greater than 8, and a statement of the level of confidence in the significance of U. (For a one-tailed test, the points of significance are found by dividing levels by two.) Since many behavioral-science data do not achieve interval scaling, the U test is a very useful alternative for one who does not wish to make the assumptions required by the parametric t test in determining if two independent samples are drawn from the same population. It is one of the most powerful of the nonparametric inferential statistics.

continued

References

1. Siegel, S., *Nonparametric Statistics for the Behavioral Sciences* (McGraw-Hill Book Co., Inc., New York, 1956).
2. Mann, H.B., and Whitney, D., "On a Test of Whether One of Two Random Variables is Stochastically Larger Than the Other," Ann. Math. Statistics *18*, 50-60 (1947).

DESCRIPTIVE TITLE	Factor Analysis (Northwestern University version originally written at Biometric Laboratory, University of Miami, Coral Gables, Fla.)
CALLING NAME	FACTOR NUCC160
INSTALLATION NAME	Vogelback Computing Center, Northwestern University
AUTHOR(S) AND AFFILIATION(S)	**Vogelback Computing Center, Northwestern University**
LANGUAGE	CDC FORTRAN IV
COMPUTER	CDC 6400
PROGRAM AVAILABILITY	Deck and listing presently available
CONTACT	Lorraine Borman, EIN Technical Representative, Vogelback Computing Center, Northwestern University, 2129 Sheridan Road Evanston, Ill. 60201 Tel.: (312) 492-3682

FUNCTIONAL ABSTRACT

FACTOR is a program that computes intercorrelations, principal components, and varimax-rotated factor loadings for a data set consisting of scored responses to a number of variables. The method employed always computes means, standard deviations, and intercorrelations at first. Then, according to user specifications, the program will either stop at that point or continue by computing the principal components and factor loadings, both normal and after varimax rotation. Another existing option permits placing squared multiple correlations in the diagonal of the correlation matrix.

DESCRIPTIVE TITLE	Student's T and F Ratio (with provision for missing data)
CALLING NAME	STDNTF2 NUCC019
INSTALLATION NAME	Vogelback Computing Center, Northwestern University
AUTHOR(S) AND AFFILIATION(S)	Vogelback Computing Center, Northwestern University
LANGUAGE	CDC FORTRAN IV
COMPUTER	CDC 6400
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Lorraine Borman, EIN Technical Representative, Vogelback Computing Center, Northwestern University, 2129 Sheridan Road, Evanston, Ill. 60201 Tel.: (312) 492-3682

FUNCTIONAL ABSTRACT

STDNTF2 is a larger, more complex version of STDNTF1 (see EIN No. 000 0013) and is capable of handling up to 500 T tests between two groups. Such a test compares an appropriate sampling (T) distribution with the normal distribution. The program tests also for consistency of performance on all tested variables.

The handling of multiple jobs, the availability of certain data transformations, and provision for cases in which the data contain missing items are features of this program.

DESCRIPTIVE TITLE	Subtle, Unbiased, Zealous Yatagen of Questionnaires	
CALLING NAME	SUZYQ NUCC150	
INSTALLATION NAME	Vogelback Computing Center, Northwestern University	
AUTHOR(S) AND AFFILIATION(S)	Program written by	Brent M. Rutherford, Northwestern University
	Converted for the CDC 6400 by	Janos B. Koplyay
LANGUAGE	CDC FORTRAN IV	
COMPUTER	CDC 6400 (Scope 3.1 O.S.)	
PROGRAM AVAILABILITY	Deck and listing presently available	
CONTACT	Lorraine Borman, EIN Technical Representative, Vogelback Computing Center, Northwestern University, 2129 Sheridan Road, Evanston, Ill. 60201 Tel.: (312) 492-3682	

FUNCTIONAL ABSTRACT

This cycling program makes possible a scoring and weighting of questionnaire items, as well as providing comprehensive item-analysis measures, test-analysis measures, and factor analysis with varimax rotation of the tetrachoric interitem correlation matrix. For the purpose of item analysis, internal as well as external criterion scores may be utilized.

The program permits the user to correct and/or weight a set of test responses (including data that have already been scored). For the cases where some data values may be skipped or missing, there is an option for the insertion of a specified average value in their place. SUZYQ provides a count of the total-response score per item and per subject, expressed in *raw* score units, *z* score units, and *t* score units. In addition, a total-test-score frequency distribution is constructed, using the same three types of units.

The *total* test score is then subjected to an item analysis with corrections for the nonindependence of the calculated correlations.

continued

Finally, a two-way variance analysis is performed, using the method of Hoyt, with basic analytic parameters computed from the elements of the summary table. Options include item analysis by criterion and a derivation of up to ten factors from the tetrachoric interitem correlation matrix (using a Kaiser varimax factor analysis).

DESCRIPTIVE TITLE	Multivariate Analysis of Variance-Covariates, Hotelling's T
CALLING NAME	MULCVR NUCC080
INSTALLATION NAME	Vogelback Computing Center, Northwestern University
AUTHOR(S) AND AFFILIATION(S)	Vogelback Computing Center, Anonymous; Northwestern University
LANGUAGE	CDC FORTRAN IV
COMPUTER	CDC 6400
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Lorraine Borman, EIN Technical Representative, Vogelback Computing Center, Northwestern University, 2129 Sheridan Road, Evanston, Ill. 60201 Tel.: (312) 492-3682

FUNCTIONAL ABSTRACT

MULCVR performs an analysis on a number (three or more) of measurements taken about an arithmetic mean (the "origin"), together with arithmetic means as specified on the Control Cards or by Hotelling's T Test.

REFERENCES

Rao, C.R., *Advanced Statistical Methods in Biometric Research* (John Wiley & Sons, Inc., New York, 1952), pp.239–248, 258–266.

Cooley, W.W., and Lohnes, P.R., *Multivariate Procedures for the Behavioral Sciences* (John Wiley & Sons, Inc., New York, 1963).

DESCRIPTIVE TITLE	FORTRAN Program for Evaluating Internal Consistency and Single-Factoredness in Sets of Multilevel Attitude Items
CALLING NAME	ATTANAL NUCC048
INSTALLATION NAME	Vogelback Computing Center, Northwestern University
AUTHOR(S) AND AFFILIATION(S)	Donald G. Morrison, Council for Intersocietal Studies, Northwestern University Donald T. Campbell, Department of Psychology, Northwestern University Leroy Wolins, Iowa State University
LANGUAGE	CDC FORTRAN IV
COMPUTER	CDC 6400
PROGRAM AVAILABILITY	Deck and listing presently available
CONTACT	Lorraine Borman, EIN Technical Representative, Vogelback Computing Center, Northwestern University, 2129 Sheridan Road, Evanston, Ill. 60201 Tel.: (312) 492-3682

FUNCTIONAL ABSTRACT

ATTANAL is a FORTRAN IV program for evaluating internal consistency and single-factoredness in sets of multilevel attitude (scale) items. The user may test the assumption that responses summarized in single scores reliably refer to a common trait, attitude, or factor being measured.

The analytical method employed avoids the Guttman unidimensionality requirement commonly associated with a reproducibility coefficient. Arbitrary or chance levels of reproducibility as well as singular tests of significance are replaced by measures of high internal consistency, reliability, and single-factoredness. Dichotomously scored items are replaced by multilevel responses, thus reducing the occurrence of irrelevant difficulty factors among the intercorrelation of items.

continued

The program operates on the data cards containing one-digit scores for each item; these are converted from response language to a scoring language to emphasize the significance of large scores. Separate analysis of pro and con items minimizes the effects of response-set and direction-of-wording effects. The polar-choice format of responses eliminates ambiguity of response effects and allows attitudes to be located on dimensions anchored by positive statements at either pole.

The program produces (for each item and the total set) the means, standard deviations, and a table of intercorrelations among items and sum. Using the Spearman—Brown Corrected Split-Half Method of Reliability, the program then produces a Kuder—Richardson coefficient of reliability based on item variances proportional to the number of items scored.

A modified factor-analysis procedure is applied to the data to obtain a test of single-factoredness. Using the intercorrelations among items, the program approximates optimal factor loadings and produces a commonality estimate (h^2) for each factor. A convergent solution is acceptable only if there are no negative factor loadings. If no convergent solution is found within 20 iterations, the program halts and prints the last iteration.

The interitem correlation matrix, the residuals matrix, as well as the mean and sigma for residuals is also reported. A Lawley chi-square test of significance is then applied to residuals, and the program produces a listing of factor loadings, the residuals matrix, and the Lawley chi-square with appropriate degrees of freedom. In addition, the program computes an index that purports to measure the relative closeness of each measure to single-factoredness; this index should approach 1.0 in single-factoredness tests.

REFERENCE

Morrison, D.G., Campbell, D.T., and Wolins, L., "A FORTRAN Program for Evaluating Consistency and Single-Factoredness in Sets of Multilevel Attitude Items," Northwestern Univ. Vogelback Computing Ctr. (1965; revised 1967).

DESCRIPTIVE TITLE	FORTRAN Program for Guttman and Other Scalogram Analyses
CALLING NAME	GUTTSCL NUCC115
INSTALLATION NAME	Vogelback Computing Center, Northwestern University
AUTHOR(S) AND AFFILIATION(S)	Roland Werner, Department of Sociology, Syracuse University Donald G. Morrison, Northwestern University
LANGUAGE	CDC FORTRAN IV
COMPUTER	CDC 6400
PROGRAM AVAILABILITY	Deck and listing presently available
CONTACT	Lorraine Borman, EIN Technical Representative, Vogelback Computing Center, Northwestern University, 2129 Sheridan Road, Evanston, Ill. 60201 Tel.: (312) 492-3682

FUNCTIONAL ABSTRACT

GUTTSCL is designed to perform Guttman and other useful analyses on data supplied in scalogram form. Originally limited to 10 items with up to 7 response categories per item and 200 respondents, the capacity has recently been expanded to 30 items and 350 persons, the Chilton and Raju indices have been added, and other minor modifications have been made. Each data set constitutes an analysis. An unlimited number of such analyses can be treated serially.

Using the Cornell technique for scalogram analysis,[1] a scalogram is displayed in the output with the frequency of the response type, the Guttman and psychometric scores of the response type, and the number of errors in the response type. The item marginal frequencies and the item marginal probabilities are also provided. This information is utilized to calculate the Guttman coefficient of reproducibility. Other coefficients of reproducibility are based upon Loevinger's,[2] Sagi's,[3] and Green's[4] methods of counting errors in a response pattern. The expected coefficients of reproducibility

continued

for each of the error-counting methods are also calculated and are based on Goodman's[5] statistical techniques.

Since a wide range of flexibility in analyzing attitude data is desirable, additional indexes are calculated. These indexes are the Loevinger homogeneity index,[2] the Kuder–Richardson formula, the corrected Kuder–Richardson formula for test reliability,[6] the Green index of consistency,[4] the Borgatta error ratio,[7] the Menzel coefficient of scalability,[8] and the Schuessler χ^2 tests[9] for the frequency distribution of response types. The source of these indexes and their method of calculation is fully described in the references cited below.

REFERENCES

1. Guttman, L., "The Cornell Technique for Scale and Intensity Analysis," Educ. Psychol. Meas. *7*, 247-280 (1947).

2. Loevinger, J., "The Technique of Homogeneous Test Compared with Some Aspects of 'Scalogram Analysis' and Factor Analysis," Psychol. Bull. *45*, 507-529 (1948).

3. Sagi, P.C., "A Statistical Test for the Significance of a Coefficient of Reproducibility," Psychometrika *24*, 19-27 (1959).

4. Green, B.F., "A Method of Scalogram Analysis Using Summary Statistics," Psychometrika *21*, 79-88 (1956).

5. Goodman, L.A., "Simple Statistical Methods for Scalogram Analysis," Psychometrika *24*, 29-43 (1959).

6. Horst, P., "Correcting the Kuder–Richardson Reliability for Dispersion of Item Difficulty," Psychol. Bull. *50*, 371-374 (1953).

7. Borgatta, E.F., "An Error Ratio for Scalogram Analysis," Public Opinion Quart. *7*, 96-100 (1955).

8. Menzel, H., "A New Coefficient for Scalogram Analysis," Public Opinion Quart. *17*, 269-280 (1953).

9. Schuessler, K.F., "A Note on Statistical Significance of Scalograms," Sociometry *24*, 312-318 (1961).

continued

Guttman, L., "The Basis for Scalogram Analysis," in *Measurement and Prediction*, S.A. Stauffer, L. Guttman, and E.A. Suchman, *et al.*, Eds. (Princeton University Press, Princeton, N.J., 1950), pp. 60-90.

Werner, R., "A FORTRAN Program for Guttman and Other Scalogram Analyses," Syracuse Univ. (CPA 257). This manual provides the program listing and the instructions for the program's use. The control of the program is achieved with control cards that allow two modes of operation: one mode provides the user with summary frequencies of the attitude data; the other, in addition, calculates various indexes appropriate to attitude data.

DESCRIPTIVE TITLE	Stepwise Multiple-Discriminant Analysis
CALLING NAME	EIDISC NUCC068
INSTALLATION NAME	Vogelback Computing Center Northwestern University
AUTHOR(S) AND AFFILIATION(S)	Donald G. Morrison, Council for Intersocietal Studies, Northwestern University Richard J. Art, Jr., Northwestern University
LANGUAGE	CDC FORTRAN IV
COMPUTER	CDC 6400
PROGRAM AVAILABILITY	Deck and listing presently available
CONTACT	Lorraine Borman, EIN Technical Representative, Vogelback Computing Center, Northwestern University, 2129 Sheridan Road, Evanston, Ill. 60201 Tel.: (312) 492-3682

FUNCTIONAL ABSTRACT

The basic stepwise procedure uses Rao's generalized distance statistic[1] to omit all variables that do not add significantly to the distance between groups. The analysis proceeds with all excluded variables set equal to zero in the discriminant functions. Group- and subject-membership probabilities are calculated. Discriminant scores are also calculated and plotted for the first two functions derived. A histogram is used if only one function is derived.

While several multiple-discriminant computer programs exist, they all are analogous to standard multiple regression in the sense that all the variables are included in the function, irrespective of their ability to increase the discriminating power of the functions derived. It is quite possible to add to the analysis variables that decrease group differences.[2] As a result, discriminant functions derived by the standard multiple-discriminant methods can often be considerably improved in their power to discriminate by the use of the stepwise procedure. There is also the added benefit of working with fewer original variables.

continued

The major question of the stepwise procedure revolves around the choice of the criteria with which to add variables. We have not used the F-ratio method of stepwise regression analysis but followed Rao[1] and used the increase in Rao's V (Ref. 1, p.257) as the entry criterion for the stepwise procedure. (The F ratio will not be able to measure the increased "distance" that a new variable will provide.) Once it is decided which variables shall be included in the analysis, then the procedure follows the methods of multiple-discriminant analysis in the reduced-variable space.

The output of the program is intended to give as much information as possible so that powerful discriminant functions can be derived. This includes measures of the discriminatory power of the derived functions, such as plots of the discriminant scores and classification of the raw input data of the discriminant function(s). Further, unclassified subjects can be classified by the program after the calculation of the discriminant function(s).

Briefly, the stepwise procedure uses the variable with the best univariate F ratio (between-groups variance to within-groups variance) as a starting point. Each of the remaining variables is tested to determine which of them adds the most to the distance statistic, a generalized Mahalanobis' D^2 described by Rao (Ref.1, p.257) and referred to as Rao's V. The greatest additional distance is tested as chi square with groups-minus-one degrees of freedom. If the "best" variable adds a significant amount, the process is repeated, testing the remaining set of variables to determine the best third variable to use in combination with the first two selected. This search-and-test process continues until it is found that none of the remaining variables adds a "significant" amount to the distance.

REFERENCES

1. Rao, C.R., *Advanced Statistical Methods in Biometric Research* (John Wiley & Sons, Inc., New York, 1952), Chaps. 7–9.

2. Morrison, D. G., and Art, R. C., Jr., "A FORTRAN Program for Stepwise Multiple Discriminant Analysis" (Northwestern University, 1967), pp.14–25. Copies may be obtained from Donald Morrison, Council for Intersocietal Studies, Northwestern University, 1818 Sheridan Road, Evanston, Ill. 60201.

DESCRIPTIVE TITLE	Multiple Regression and Correlation Analysis, Modified with Plots
CALLING NAME	PLOTYY NUCC079
INSTALLATION NAME	Vogelback Computing Center, Northwestern University
AUTHOR(S) AND AFFILIATION(S)	Vogelback Computing Center
LANGUAGE	CDC FORTRAN IV
COMPUTER	CDC 6400
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Lorraine Borman, EIN Technical Representative, Vogelback Computing Center, Northwestern University, 2129 Sheridan Road, Evanston, Ill. 60201 Tel.: (312) 492-3682

FUNCTIONAL ABSTRACT

This program calculates basically the same information as does BMD29 (EIN No. 000 0012). However, it also plots the predicted Y versus observed Y values of the dependent variable.

As in BMD29, the independent variables are listed in the order of their importance, based on the reduction of sums of squares of the dependent variable attributable successively to each independent variable. The maximum number of variables that can be processed by this program is 59.

DESCRIPTIVE TITLE	Multiple Regression and Correlation Analysis
CALLING NAME	BMD29 NUCC043
INSTALLATION NAME	Vogelback Computing Center, Northwestern University
AUTHOR(S) AND AFFILIATION(S)	Program originated at UCLA Medical Center as part of the Biomedical Statistics Package
LANGUAGE	CDC FORTRAN IV
COMPUTER	CDC 6400
PROGRAM AVAILABILITY	Deck and listing presently available
CONTACT	Lorraine Borman, EIN Technical Representative, Vogelback Computing Center, Northwestern University, 2129 Sheridan Road, Evanston, Ill. 60201 Tel.: (312) 492-3682

FUNCTIONAL ABSTRACT

This program is a modification and extension of BMD06 of the UCLA biomedical package. The primary modification is that the independent variables are listed in the order of their importance, based on the reduction of sum of squares of the dependent variable attributable successively to each independent variable. Additional output includes cumulative

(1) standard errors of estimate,
(2) sums of squares,
(3) proportions of variance,
(4) F values,
(5) multiple correlation coefficients.

The maximum number of variables that can be processed by this pro- is 59.

The upper limitation on sample size is 99,999 and the lower limitation on sample size is two greater than the number of variables.

continued

This program can perform a different transformation of any variable and generate new variables if desired, according to the codes specified in the transgeneration cards or through lagging. Any number of variables can be generated, but the total number of variables must not exceed 59.

Any variable, original or generated, can be named the dependent variable. There is no limit to the number of *replacements* (see Replacement—Deletion Cards).

The maximum number of variables that can be deleted at one time is 32. However, there is no limit to the number of deletions of different sets of variables.

The standard format (12F6.0) for input data or a variable format may be used.

This program uses a reduction of the sum-of-the-squares method to calculate regression and correlation coefficients to fit data points.

REFERENCES

Although no specific references are provided, the user is directed to check statistical references for the Von Neuman ratio and the Durbin—Watson coefficient of correlation.

DESCRIPTIVE TITLE	Student's T Test
CALLING NAME	STDNTF1 NUCC018
INSTALLATION NAME	Vogelback Computing Center, Northwestern University
AUTHOR(S) AND AFFILIATION(S)	Vogelback Computing Center, Northwestern University
LANGUAGE	CDC FORTRAN IV
COMPUTER	CDC 6400
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Lorraine Borman, EIN Technical Representative, Vogelback Computing Center, Northwestern University, 2129 Sheridan Road, Evanston, Ill. 60201 Tel.: (312) 492-3682

FUNCTIONAL ABSTRACT

STDNTF1 computes 50 T tests between two groups. Such tests compare an appropriate sampling (T) distribution with the normal distribution. The program also tests for consistency of performance on all variables within the test. Also, the handling of multiple jobs is possible. (See STDNTF2, EIN No. 000 0003.)

DESCRIPTIVE TITLE	Northwestern University Correlation Analysis, Version 1
CALLING NAME	NUCORR1 NUCC070
INSTALLATION NAME	Vogelback Computing Center, Northwestern University
AUTHOR(S) AND AFFILIATION(S)	Vogelback Computing Center, Northwestern University
LANGUAGE	CDC FORTRAN IV
COMPUTER	CDC 6400
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Lorraine Borman, EIN Technical Representative, Vogelback Computing Center, Northwestern University, 2129 Sheridan Road, Evanston, Ill. 60201 Tel.: (312) 492-3682

FUNCTIONAL ABSTRACT

NUCORR1 may be used to compute sums, sums of squares, crossproducts, means, standard deviations, and all or any desired section of the correlation matrix. There is also the capability of classifying data into groups on the basis of a single property.

NUCORR1 uses a maximum-likelihood estimator that deals with the principle of point estimates. This method is quite consistent and, for large numbers of samples, very efficient and hightly accurate. The basis concept of point estimation is the consideration of every possible value of the parameter and the calculation of the probability that that particular sample would occur in the normal distribution if it were the true value of the parameter. NUCORR1 also serves as an asymptotic estimator in that, as the size of the sample increases, the parameter will approach a limit (defined by the preceding values).

DESCRIPTIVE TITLE	Automatic Interaction Detector
CALLING NAME	AID NUCC113
INSTALLATION NAME	Vogelback Computing Center, Northwestern University
AUTHOR(S) AND AFFILIATION(S)	J. Sonquist The University of Michigan
LANGUAGE	CDC FORTRAN IV
COMPUTER	CDC 6400
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Lorraine Borman, EIN Technical Representative, Vogelback Computing Center, Northwestern University, 2129 Sheridan Road Evanston, Ill. 60201 Tel.: (312) 492-3682

FUNCTIONAL ABSTRACT

AID was originally programmed for the IBM 7090 in the MAD language at The University of Michigan. A detailed discussion of the theory, methods, and control parameters of the program are contained in Ref. 1. Since 1964, two parameters have been added to the main-parameter card.

AID is focused on a particular kind of data-analysis problem, characteristic of many social-science research situations, in which the purpose of the analysis involves more than the reporting of descriptive statistics but may not necessarily involve the exact testing of specific hypotheses. In this type of situations the problem is often one of determining which of the variables, for which data have been collected, are related to the phenonmenon in question, under what conditions, and through what intervening processes, with appropriate controls for spuriousness.

AID is useful in studying the interrelationships among a set of up to 37 variables. Regarding one of the variables as a dependent variable, the analysis employs a nonsymmetrical branching process, based on variance-analysis techniques, to subdivide the sample into a series of subgroups that maximizes one's ability to predict values of the dependent variable. Linearity and additivity

continued

assumptions inherent in conventional multiple-regression techniques are not required. AID will handle variables that are only nominal scales, i.e., mere classifications.

REFERENCES

1. Morgan, J.N., and Sonquist, J.A., "The Detection of Interaction Effects: A Report on a Computer Program for the Selection of Optimal Combinations of Explanatory Variables," Univ. Mich. Inst. Soc. Res. Survey Res. Ctr. Monograph No. 35 (1964); copies may be obtained from the EIN office for the cost of duplication and mailing.

DESCRIPTIVE TITLE	Generalized Stepwise Regression
CALLING NAME	B34T NUCC111
INSTALLATION NAME	Vogelback Computing Center, Northwestern University
AUTHOR(S) AND AFFILIATION(S)	This program is a single-precision version of the UCLA BIMED 34, rewritten by Hodson Thornber, The University of Chicago, in FORTRAN II and converted for the CDC 6400 by Donald Morrison, Northwestern University.
LANGUAGE	CDC FORTRAN IV
COMPUTER	CDC 6400
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Lorraine Borman, EIN Technical Representative, Vogelback Computing Center, Northwestern University, 2129 Sheridan Road, Evanston, Ill. 60201 Tel.: (312) 492-3682

FUNCTIONAL ABSTRACT

B34T calculates multiple linear regression in a stepwise manner, inserting at each step that variable which is eligible and has the highest partial correlation with the dependent variable, given that the previously selected variables are in the equation. The eligibility of a variable to be entered is specified by the use of Order Cards and Forcing Mode Cards. The latter, if present, delineate a set of variables to be forced in, a set to be forced out, and a set to be entered stepwise by the partial correlation criterion mentioned above. The order card specifies the order in which a specified set of variables is entered, as follows. B34T forgets completely about variables forced out. It then enters all variables listed on Order Card(s), holding in the equation those forced in, regardless of how low their partial correlations may become. If any forced-in variables remain after the ordered variables have been entered, they are taken as a group and entered by the correlation criterion. After all ordered and all forced variables are in, any unforced variables are entered stepwise.

continued

REFERENCES

Thornber, H., "Manual for (B34T, 8 Mar. 1966) A Stepwise Regression Program," Northwestern Univ. Rep. 6603, a general algebraic description of the standard regression model (with and without the constant), main features and operations of the program, use of special features of the program, and a macro flow for program logic.

"Writeup for (B34T, 8 Mar. 1966) A Stepwise Regression Program" gives detailed instructions for filling out individual Control Cards. This abbreviated writeup may be used by casual users who wish to perform simple and multiple regressions with an expanded set of transgenerations that includes lagging of the variables.

DESCRIPTIVE TITLE	Biserial—Point-Biserial Correlation Program
CALLING NAME	BISR NUCC117
INSTALLATION NAME	Vogelback Computing Center, Northwestern University
AUTHOR(S) AND AFFILIATION(S)	N. John Castellan, Jr. Vogelback Computing Center Stephen W. Link Vogelback Computing Center
LANGUAGE	CDC FORTRAN IV
COMPUTER	CDC 6400
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Lorraine Borman, EIN Technical Representative, Vogelback Computing Center, Northwestern University, 2129 Sheridan Road, Evanston, Ill. 60201 Tel.: (312) 492-3682

FUNCTIONAL ABSTRACT

BISR has been designed to compute either biserial or point-biserial correlation coefficients. At the user's option, the biserial coefficients will be computed, assuming that a particular discrete item constitutes a subscore for the continuous variable. (This case is analogous to an item-total correlation.) The following data for each discrete variable are obtained from this analysis.

(a) means of continuous variables for failing groups

(b) means of continuous variables for passing groups

(c) means for the continuous variables

(d) proportion of cases in the passing group

(e) standard deviation of continuous variables for those cases with a pass or fail response on the associated discrete variable

(f) correlation coefficients

continued

Computational Procedure

Formulas Used:

$$r_b = (M_2 - M_1)\ P_1\ \ P_2/z\ s_y,$$

$$r_{pb} = [(M_2 - M_y)/s_y]\ (P_2/P_1)^{1/2}.$$

If $r_b{}^1$ is to represent an item-total correlation,

$$r_b{}^1 = [r_b\ s_y - (P_1\ P_2\ /z)]/\ (s_y{}^2 + P_1\ P_2 - 2r_b\ s_y\ z)^{1/2},$$

where P_1 is the proportion of cases in the failing group (0's) and P_2 is the proportion of cases in the passing group (1's or 2's); M_1 and M_2 are the means of the continuous variable for failing and passing groups, respectively, and M_y is the mean for the continuous variable; s_y is the standard deviation for the continuous variable for those cases with a pass or fail response on the associated discrete variable; and z is the ordinate of the unit normal curve at the point where P_1 (or P_2) cases are cut off.

$$s_y{}^2 = [N\ \Sigma\ X^2 - (\Sigma\ X)^2]/\ N(N - 1)\ .$$

REFERENCES

Castellan, N. J., Jr., and Link, S. W., " Biserial Point-Biserial Correlation Program for IBM 709/7090," Behavioral Sci. *8*, 368–369 (Oct. 1963).

Henrysson, S., "Correction of Item-Total Correlations in Item Analysis," Psychometrika *28*, 211–218 (1963).

DESCRIPTIVE TITLE	Canonical Correlation
CALLING NAME	CANON NUCC168
INSTALLATION NAME	Vogelback Computing Center, Northwestern University
AUTHOR(S) AND AFFILIATION(S)	University of Miami Biometric Laboratory
LANGUAGE	CDC FORTRAN IV
COMPUTER	CDC 6400
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Lorraine Borman, EIN Technical Representative, Vogelback Computing Center, Northwestern University, 2129 Sheridan Road, Evanston, Ill. 60201 Tel.: (312) 492-3682

FUNCTIONAL ABSTRACT

The program computes canonical correlations between two sets of variables and tests their significance. It also scores each subject, using the weights associated with the canonical correlations. The methods used in these calculations may be found in the references cited below.

REFERENCES

Hotelling, H., "Relations between Two Sets of Variates," Biometrika *28*, 321–377 (1936).

Bartlett, M. S., "The Statistical Significance of Canonical Correlations," Biometrika *32*, 29–37 (1941).

DESCRIPTIVE TITLE	χ^2 and C Statistics
CALLING NAME	CHISQR NUCC028
INSTALLATION NAME	Vogelback Computing Center Northwestern University
AUTHOR(S) AND AFFILIATION(S)	Vogelback Computing Center Northwestern University
LANGUAGE	CDC FORTRAN IV
COMPUTER	CDC 6400
PROGRAM AVAILABILITY	Decks and listings presentaly available
CONTACT	Lorraine Borman, EIN Technical Representative, Vogelback Computing Center, Northwestern University, 2129 Sheridan Road Evanston, Ill. 60201 Tel.: (312) 492-3682

FUNCTIONAL ABSTRACT

CHISQR forms a contingency table from numeric data and calculates χ^2, C, the percentages of samples in the categories, and the percentages of the categories in the samples.

CHISQR reads in *numeric* data for as many as 10 samples, each consisting of as many as 1000 subjects. The range of sample values for all samples can be divided into as many as 10 subsets, called categories; a contingency table is then formed by determining the frequencies with which data from ISMPLS of the samples ($2 \le \text{ISMPLS} \le 10$) fall into the categories. χ^2 and C (the contingency coefficient) are printed out, the table is displayed, and the percentages of samples in the categories and categories in the samples are shown. Alternatively, the program will accept input of the actual frequencies and/or expected frequencies instead of calculating them from raw data and frequencies, respectively.

Computational Procedure

χ^2 Test

1. The *relative frequencies* with which m *independent* samples (rows) are classified into n mutually exclusive categories (columns) are calculated and treated as interval measurements. This

continued

implies only that the raw data must have *nominal* (classificatory) measurement. A table of these frequencies is called a *contingency table*.

2. H_0 indicates that there is no difference between the m samples in the proportion of observations that fall into the n categories—i.e., that the rows of frequencies have come from the same or identical populations. H_1 must be two-tailed if $m > 2$, $n > 2$ (i.e., $df > 1$), as explained below.

3. χ^2 is calculated from the contingency table; its sampling distribution is approximated by the chi-squared distribution with $df[=(m-1)*(n-1)]$ degrees of freedom.

The χ^2 test cannot be applied if the expected frequencies are too small. If $df = 1$, a Yates-corrected χ^2 is calculated. Difficulty with small expected frequencies can be avoided by increasing the number of degrees of freedom or by combining "similar" categories (columns) of the contingency table. The χ^2 test is insensitive to the effects of order when $df > 1$.

4. The power efficiency of the χ^2 test is difficult to determine, since it is applied usually to data that cannot be analyzed by any alternative test. However, its limiting power is one as the number of degrees of freedom increases.

Contingency Coefficient

1. The contingency coefficient (C) is computed from the χ^2 contingency table and will have no correlation to the population when H_0 exists (i.e., the value of C arose from chance). When H_1 exists, the samples for C were correlated from a contingency table where $df = [(m-1)*(n-1)] > 1$. This is an indicator that C is significant and accurate.

2. When $C = 0$, there is a complete lack of association between the variables being measured.

3. When $C \neq 1$ $(0 < C < 1)$, there is a complete dependence between the variables.

4. When C approaches 1, the degrees of freedom approach infinity.

5. Two contingency coefficients can not be compared unless they were derived from the *same size tables*. These coefficients are not directly comparable to any other measure of correlation.

DESCRIPTIVE TITLE	Discriminant Analysis for Two Groups
CALLING NAME	DISCRIM NUCC169
INSTALLATION NAME	Vogelback Computing Center Northwestern University
AUTHOR(S) AND AFFILIATION(S)	Biometrics Laboratory University of Miami
LANGUAGE	CDC FORTRAN IV
COMPUTER	CDC 6400
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Lorraine Borman, EIN Technical Representative, Vogelback Computing Center, Northwestern University, 2129 Sheridan Road Evanston, Ill. 60201 Tel.: (312) 492-3682

FUNCTIONAL ABSTRACT

DISCRIM computes a discriminant function for contrasting two groups and tests its significance. It also scores each subject by using the discriminant-function coefficients. The analysis may be repeated on subsets of the original variables.

Rao's[1] discriminant-function coefficients are multiplied by a constant so that the sum of squares of the coefficients is equal to one; these are printed under the heading RAW DISCRIMINANT FUNCTION COEFFICIENTS.

REFERENCES

1. Rao, C.R., *Advanced Statistical Methods in Biometric Research* (John Wiley & Sons, Inc., New York, 1952), pp. 237–238, 246–248.

Anderson, T.W., *An Introduction to Multivariate Statistical Analysis* (John Wiley & Sons, Inc., New York, 1958), pp. 126–138.

DESCRIPTIVE TITLE	Eigensystems of Nonsymmetric, Real Square Matrices
CALLING NAME	EIGSYS NUCC126
INSTALLATION NAME	Vogelback Computing Center, Northwestern University
AUTHOR(S) AND AFFILIATION(S)	James Van Ness Department of Electrical Engineering, Northwestern University
LANGUAGE	CDC FORTRAN IV
COMPUTER	CDC 6400
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Lorraine Borman, EIN Technical Representative, Vogelback Computing Center, Northwestern University, 2129 Sheridan Road, Evanston, Ill. 60201. Tel.: (312) 492-3682

FUNCTIONAL ABSTRACT

EIGSYS produces all eigenvalues and eigenvectors (both real and complex) of either symmetric or nonsymmetric square matrices having real elements. EIGSYS calls various subroutines to calculate the eigenvalues, the determinant and the eigenvectors. The trace and determinant of input matrix A is compared with the sum and product, respectively, of the E values. The righthand eigenvectors X(I) and the lefthand eigenvectors V(J) of A are printed along with their respective residual vectors (A*X(I) - LAMBDA(I)*X(I)) and V(J)*A-V(J)*LAMBDA(J)). The squares of the residual vectors are printed for convenience in checking the accuracy of the eigensystem. After finding each E vector, EIGSYS then uses the Rayleigh quotient[1] scheme to correct the original E value, which in turn produces a corrected E vector along with new residuals. A second correction is applied that produces a further refinement of the eigenvalue only. The sum and product of the corrected E values are printed at the end of the output for each matrix. EIGSYS repeats this computation for each matrix in the data-input stream.

REFERENCES

1. Martin, R.S., and Wilkinson, J.H., "Solution of Symmetric and Asymmetric Band Equations and the Calculation of Eigenvectors of Band Matrices," Numer. Math. *9*, 279–301 (1959).

DESCRIPTIVE TITLE	Multivariate Analysis of Variance
CALLING NAME	MANOVA NUCC170
INSTALLATION NAME	Vogelback Computing Center, Northwestern University
AUTHOR(S) AND AFFILIATION(S)	Biometrics Laboratory, University of Miami
LANGUAGE	CDC FORTRAN IV
COMPUTER	CDC 6400
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Lorraine Borman, EIN Technical Representative, Vogelback Computing Center, Northwestern University, 2129 Sheridan Road, Evanston, Ill. 60201 Tel.: (312) 492-3682

FUNCTIONAL ABSTRACT

MANOVA performs univariate and multivariate analyses of variance, of covariance, and of regression. It provides an exact solution in either the orthogonal or nonorthogonal case. Options include single- or multiple-degree-of freedom contrasts in the main effects or interactions, transformations of variables, and orthogonal polynomial contrasts with equally or unequally spaced points. Reanalyses may be done with different criteria, covariates, contrasts, and models.

REFERENCES

Bock, R.D., "Programming Univariate and Multivariate Analysis of Variance," Technometrics *5*, 95–117 (1963).

Graybill, F.A., *An Introduction to Linear Statistical Models* (McGraw-Hill Book Co., Inc., New York, 1961), Vol. 1, pp. 223--253.

Rao, C.R., *Advanced Statistical Methods in Biometric Research* (John Wiley & Sons, Inc., New York, 1952), pp. 236–272.

Roy, S.N., *Some Aspects of Multivariate Analysis* (John Wiley & Sons, Inc., New York, 1957), pp. 76–86.

DESCRIPTIVE TITLE	95X95 Factor Analysis with Varimax Rotation
CALLING NAME	MESA1 NUCC016
INSTALLATION NAME	Vogelback Computing Center, Northwestern University
AUTHOR(S) AND AFFILIATION(S)	Vogelback Computing Center, Northwestern University
LANGUAGE	CDC FORTRAN IV
COMPUTER	CDC 6400
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Lorrain Borman, EIN Technical Representative, Vogelback Computing Center, Northwestern University, 2129 Sheridan Road Evanston, Ill. 60201 Tel.: (312) 492-3682

FUNCTIONAL ABSTRACT

MESA1 can correlate, factor, and rotate a maximum of 95 variables. The maximum sample size that can be used with the program is 32 768 observations. Sample size should be larger than number of variables.

MESA1 is composed of three parts: a main program, an eigenvalue subroutine, and a varimax-rotation subroutine. The eigenvalues are calculated by a modified Jacobian method that closely parallels the routine given by Greenstadt in the Ralston and Wilf volume.[1] The varimax subroutine is from BIMD17 package and is based on the Kaiser algorithm.[2] The method of principal-components orthogonal rotation is used.

All of the output of the program is in the form of naturally arranged tables. The variables in the tables can be labeled by name as well as by number. The output of the program can include

—means, standard deviations, the third and fourth moments of each variable

—the standard errors of each of the above

continued

—triangular matrix of product moment correlation coefficients, with communality estimates in the main diagonal

—a table of positive eigenvalues, with percentages and cumulative percentages based upon both the total variance of the matrix and the variance accounted for by the factors to be rotated

—matrix of rotated factors

—matrix of unrotated factors. (The number of factors included in this and subsequently listed output is subject to the restrictions explained in the following section.)

The program has the following options available, which are determined by Control-Card entries.

A. Input and Communality Options

The input to the program can be either a series of $N (N < 32\,768)$ observations of M variables each $(M \leq 95)$ or a MXM correlation matrix. If the input is in the form of observations, the following communality options can be computed for each variable X_i

1. All 1's
2. Squared multiple correlation of X_i with the remaining X_j, $i \neq j$
3. Absolute value of the largest correlation of X_i with X_j, $i \neq j$

If a matrix input is used, only the following options are avaiable for communality estimate for variables X_i

4. All 1's
5. Absolute value of largest correlation of X_i with X_j
6. The diagonal element r_{ii} *in the data input matrix*

No factor scores are available with options 4, 5, and 6. In these cases, Cols. 36—37 of the Problem Card *must* be blank.

B. Rotation Options

1. Rotation is optional and can be suppressed
2. If rotation is not suppressed, the following series of conditions and options determines the number of variables that will be rotated. (These conditions also determine the number of factors included in the factor matrix.)

The following conditions always hold

a. A maximum of *19 factors* can be rotated

b. Only factors with *positive eigenvalues* can be rotated

continued

c. A maximum of 100% of the communality of the original matrix can be accounted for by factors. (In practice, this condition means that, if the first k largest factors account for 100% of the original variance, then *no addition* factors will be calculated.)

The largest number of factors that can be rotated is the minimum of the three values determined by the above conditions. The number of factors to be rotated can be *further* restricted by the following options.

d. Specification of a *maximum number of factors*

e. Specification of the *minimum eigenvalue* to be included among rotated factors. (This specification can be a constant, or can be evaluated as the absolute value of the largest negative eigenvalue)

f. Specification of a *minimum/maximum-factor loading* for factors to be included in the rotation. (In practice, this option means that it is possible to exclude from rotation any factors that have no loadings at or above the specified value. This option should be used with caution because the program will stop calculating factors at the *first* factor that fails this maximum-factor-loading criterion. It sometimes happens that a factor will fail this min/max test when a factor with a smaller eigenvalue would not fail it.)

If any of the communality options d, e, or f are used, the minimum value, so determined, serves as the limit on the number factors to be rotated, if that value is less than the value determined by condition a, b, and c. The options d, e, and f can be suppressed by leaving the appropriate Control-Card columns blank.

REFERENCES

1. Greenstadt, J., "The Determination of the Characteristic Roots of a Matrix by the Jacobi Method," in *Mathematical Methods for Digital Computers*, Ralston and Wilf, Eds. (John Wiley & Sons, Inc., New York, 1959), Chap. 7.

2. Kaiser, H.F., "The Varimax Criterion for Analytical Rotation in Factor Analysis," Psychometrika *23*, No. 3 (Sept. 1958).

DESCRIPTIVE TITLE	Northwestern University Cross Classification and Tabulation
CALLING NAME	NUCROS NUCC017
INSTALLATION NAME	Vogelback Computing Center, Northwestern University
AUTHOR(S) AND AFFILIATION(S)	Betty Benson Vogelback Computing Center, Northwestern University
LANGUAGE	CDC FORTRAN IV
COMPUTER	CDC 6400
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Lorraine Borman, EIN Technical Representative, Vogelback Computing Center, Northwestern University, 2129 Sheridan Road, Evanston, Ill. 60201 Tel.: (312) 492-3682

FUNCTIONAL ABSTRACT

NUCROS is a general program for preparing crossclassifications (also called *crosstabulations* or *contingency tables*) in two, three, or four dimensions. For all tables, it is also possible to obtain row and column percentages, chi square, Kendall's tau, the Goodman—Kruskal gamma, and Somers D. A maximum of 99 problems may be executed in one run; each run is able to produce up to 72 tables from a given set of Control Cards and data. Data input may be from cards or tape. For each problem, a maximum of 9999 cases, with up to 80 variables/case, may be processed. The input data must be of the integer type; however, they may be recoded.

REFERENCES

Janda, K., *Data Processing* (Northwestern University Press, Evanston, Ill., 1968), 2nd ed.

Bonato, R. R., and Waxman, B. D., "A General Cross-Classification Program for Digital Computers," Behavioral Sci. *6*, No. 4, 347—357 (Oct. 1961).

DESCRIPTIVE TITLE	Critical-Path Summary—Probability Statistic Based on Normal Curve
CALLING NAME	PERTC NUCC022
INSTALLATION NAME	Vogelback Computing Center Northwestern University
AUTHOR(S) AND AFFILIATION(S)	Vogelback Computing Center Northwestern University
LANGUAGE	CDC FORTRAN
COMPUTER	CDC 6400
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Lorraine Borman, EIN Technical Representative, Vogelback Computing Center, Northwestern University, 2129 Sheridan Road, Evanston, Ill. 60201 Tel.: (312) 492-3682

FUNCTIONAL ABSTRACT

PERTC produces a detailed scheduling network, final critical-path summary, and probability statistic for events that are normally distributed about some mean.

REFERENCES

Rogers, H.B., "Construction Project Pert Network," Northwestern Univ. Grad. School Business (Jan. 1964; unpublished). A reference copy may be obtained from the EIN office for the cost of duplicating and mailing.

DESCRIPTIVE TITLE	Punch and Print Rankings
CALLING NAME	PPRANK NUCC023
INSTALLATION NAME	Vogelback Computing Center, Northwestern University
AUTHOR(S) AND AFFILIATION(S)	Vogelback Computing Center, Northwestern University
LANGUAGE	CDC FORTRAN IV
COMPUTER	CDC 6400
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Lorraine Borman, EIN Technical Representative, Vogelback Computing Center, Northwestern University, 2129 Sheridan Road, Evanston, Ill. 60201 Tel.: (312) 492-3682

FUNCTIONAL ABSTRACT

PPRANK is able to rank raw data for as many as 100 subjects on as many as 140 variables and to punch out these ranks. A nonparametric Spearman product—moment correlation is performed upon the ranks, with respect to the raw data. Consensus ratings may be produced and tested for significance by using the Kendall W coefficient of concordance. No missing-data option is available.

The subjects are ranked with respect to each variable. The rank 1 is assigned to the subject with the smallest (i.e., most negative) sample value of the variable being considered. Subjects that tie at a sample value are assigned a rank that is the average of the tied ranks. A consensus ranking of the subject can be selected by specifying as many as 20 variables as the ones whose ranks are to be summed and ranked. The consensus ranking itself then becomes an additional variable. Ties among the sums of ranks are resolved by comparison of the sums of squares of ranks; the smaller the sum of squares of ranks, the lower the rank assigned the sum of ranks. Should ties still persist, an average rank is assigned. It is not statistically sound to form consensus ranks from variables that are not in accord (as measured by a W test at the significance level set by the user).

DESCRIPTIVE TITLE	Rank-Order Statistics
CALLING NAME	RKSTAT NUCC020
INSTALLATION NAME	Vogelback Computing Center Northwestern University
AUTHOR(S) AND AFFILIATION(S)	Vogelback Computing Center Northwestern University
LANGUAGE	CDC FORTRAN
COMPUTER	CDC 6400
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Lorraine Borman, EIN Technical Representative, Vogelback Computing Center, Northwestern University, 2129 Sheridan Road, Evanston, Ill. 60201 Tel.: (312) 492-3682

FUNCTIONAL ABSTRACT

RKSTAT ranks as many as 100 subjects on as many as 80 variables and calculates the following rank-order statistics that compare the ranks assigned to two or more variables: Wilcoxon—Mann—Whitney U, Kruskal—Wallis H, Spearman rank correlation coefficient RS, Kendall coefficient of concordance W. Sample values for these statistics are printed out with directions for the testing of their significance.

The statistical comparison tests in this category assume that the raw data have only the strength of *ordinal* measurement. A raw-data observation is assigned a *rank* that is given by one plus the number of observations that are numerically smaller than it (if IRANK = 0) or numerically larger than it (if IRANK = 1). In calculating the statistics in these tests, the ranks are treated as interval measurements without implying that the raw data have the strength of an interval scale. These statistics universally assume *underlying continuity*; therefore, the probability of a tie in the raw data is zero. When a tie does occur, however, it is immediately assumed that the tie is an illusion produced by inaccurate measurement and each tied observation is assigned the average of the ranks available to the tied observations. For example, if a single observation of 1.53 has been assigned the rank 9 and four observations are tied at 1.54, each of the four tied observations

continued

will be assigned the rank 11.5 and any single observation occurring at 1.55 would get the rank 14.

Wilcoxon—Mann—Whitney U Statistic Test—compares two, independent samples of equal or unequal size. The U test is more sensitive to differences in location (central tendency) than to differences in dispersion or skewness. The Kolmogorov—Smirnov D test is more suitable than the U test if one wishes to test for *all* kinds of differences.

For fixed, small α, the power efficiency of the U test, as compared with the t test, approaches 95.5% (i.e., $3/\pi$ X 100) as N approaches infinity. This shows that the U test is very powerful for large samples and even moderately sized samples.

Kruskal—Wallis H Statistic Test—compares k independent samples of equal or unequal length. Like the U test, the H test is primarily a test of location.

For fixed small α, the H test, as compared with the F test, has an asymptotic efficiency of 95.5% (i.e., $3/\pi$ X 100).

Spearman Rank Correlation Coefficient RS—gives the correlation between two matched samples (of equal length N).

The Spearman RS has approximately 91% power efficiency when compared with the Pearson r (i.e., when the matched data are, in fact, taken from a bivariate normal distribution, are measured on an interval scale, and are correlated in the population, then r will detect this correlation at the significance level with 91 observations for every 100 observations that are required to detect this correlation at this significance level by RS).

Kendall Coefficient of Concordance W—gives the correlation between k matched samples of equal length N. *Note:* a significant W does not imply that the rankings are objective or "right," but only "consensual." I.e., each sample appears to have its observations ranked on the same (possible erroneous) criteria.

The power of W is 91% as compared to the average of the pairwise Pearson r correlations of the k samples.

Interested persons are directed to the source listed in the Reference for a fuller mathematical explanation of the various tests.

REFERENCE

Siegel, S., *Nonparametric Statistics* (McGraw—Hill Book Co., Inc., New York, 1956).

DESCRIPTIVE TITLE	Interrupt Time Series—Three Tests of Significance
CALLING NAME	TIMEX NUCC049
INSTALLATION NAME	Vogelback Computing Center, Northwestern University
AUTHOR(S) AND AFFILIATION(S)	Joyce Sween, Department of Sociology, Northwestern University
LANGUAGE	CDC FORTRAN IV
COMPUTER	CDC 6400
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Lorraine Borman, EIN Technical Representative, Vogelback Computing Center, Northwestern University, 2129 Sheridan Road, Evanston, Ill. 60201 Tel.: (312) 492-3682

FUNCTIONAL ABSTRACT

TIMEX performs three statistical tests on data of the interrupted time-series form: (1) the Mood test, (2) the Walker—Lev covariance tests, and (3) double extrapolation.

In addition to the T or F values of each statistical test of significance, the following information may be obtained (options to be indicated on card input).

Mood Test

Standard error for t test
Estimated y value for the first postchange point
Obtained y value for the first postchange point
Slope and intercept for the prechange points
Variance of the obtained prechange points from the regression estimated values

Walker—Lev Tests

Test 1

Numerator and denominator sum of squares for F
Regression estimates (slope and intercept) for prechange values

continued

Estimated y value at experimental change (i.e., midway between last pre- and first post-change points), based upon prechange regression estimates
Regression estimates for postchange values
Estimated y value at experimental change, based upon postchange regression estimates
Common within-groups regression estimates for the pre- and post-change values
Estimated y value at experimental change, based upon each of the above regression estimates

Test 2

Numerator and denominator sum of squares for F

Test 3

Numerator and denominator sum of squares for F
Regression estimates (slope and intercept) for the total series
Estimated y value at experimental change, based upon the above regression estimates

Double-Extrapolation Technique for Mood Test

Standard error for T test
Regression estimates for pre- and post-change values
Variance of the obtained pre- and post-change points from the values predicted on the basis of the above regression estimates
Estimated y value at experimental change based upon pre- and post-change regression estimates

A data plot of the time series may also be obtained.

REFERENCES

Sween, J., and Campbell, D.T., "A Study of the Effect of Proximally Autocorrelated Error on Tests of Significance for the Interrupted Time Series Quasi-Experimental Design," Northwestern Univ. (Aug. 1965; unpublished).

Sween, J., and Campbell, D.T., "The Interrupted Time Series as Quasi-Experiment: Through Tests of Significance," Northwestern Univ. (Aug. 1965; unpublished).

DESCRIPTIVE TITLE	Test Scorer and Statistical Analysis
CALLING NAME	TSSA NUCC072
INSTALLATION NAME	Vogelback Computing Center Northwestern University
AUTHOR(S) AND AFFILIATION(S)	Richard Wolf and Leopold Klopfer The University of Chicago K. Jones Harvard University Graduate School of Education A. Gasche, B. Wright, and C. Bradford The University of Chicago
LANGUAGE	CDC FORTRAN IV
COMPUTER	CDC 6400
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Lorraine Borman, EIN Technical Representative, Vogelback Computing Center, Northwestern University, 2129 Sheridan Road, Evanston, Ill. 60201 Tel.: (312) 492-3682

FUNCTIONAL ABSTRACT

TSSA scores multiple-response tests that have a single correct response for each item, computes test and item statistics, computes the tetrachoric interitem correlation matrix, and performs a factor analysis and varimax rotation. The point-biserial correlation of each item with a criterion score also may be obtained. The number of alternative responses may vary from item to item.

Multiple-scoring keys may be used with the same set of data cards. This feature makes TSSA applicable for scoring and analysis of the Kuder preference record and similarly constructed multikeyed instruments. In addition, multiple jobs may be run.

Computational results include the following.

Individual Scores, including raw scores and scores corrected for guessing

continued

Item-Response Information, including the proportion of subjects selecting the correct response for each item (i.e., difficulty) and the proportion of subjects selecting each response

Test Statistics,[1] including the mean, standard deviation, skewness, and kurtosis (and their standard errors computed from formulas outlined by Fisher); test reliability (computed by Kuder–Richardson formula 20); a validity coefficient; and a Pearson product–moment correlation

Item-Analysis Information,[2,3] including the proportion of subjects passing an item, the item standard deviation, point-biserial correlations, reliability index, and validity index (computed from formulas given by Guilford and Gullicksen)

Tetrachoric Interitem Correlations[4]

Factor Analysis of Interitem Correlation Matrix,[5] with varimax rotation and plot

Multiple-Scoring Keys

A special feature of TSSA is the use of Multiple-Scoring Keys. Such a feature has two uses. (1) It allows a single set of responses to be scored in more than one way; practical application of this is the use of the program with data from an instrument such as the Kuder preference record where several scores, based on analyses of the same set of items, are desired. (2) A more common situation is the use of the program to obtain subtest scores and a total test score from one instrument. An application of this might be to score all the responses on a reading test and to obtain scores for reading speed, reading vocabulary, and level of comprehension, as well as a total reading score.

Multiple Jobs

The limitation of the procedures in the analyses of subtests is that only the raw scores for each individual or each subject will be meaningful. Corrected scores will *not* have any meaning. Thus, if an analysis excludes any items initially read, no corrected score will be printed. The reason for this is that the program treats *all* items excluded from analysis as incorrect responses. Thus, the use of a formula for correcting scores for guessing is fallacious in this instance. If the user desires meaningful corrected scores for subtests, he will have to reproduce his Data Cards and run multiple jobs, selecting for reading on any one job only those items for which analysis is desired, i.e., the subtest.

continued

The use of the Multiple-Jobs feature for scoring subtests yields meaningful corrected scores, as already mentioned. However, the advantage of the Multiple-Scoring-Key feature of the program is that the data need be read from cards only once. Subsequent reading of the data is done from binary tape (tape B3), thus achieving savings of time and, hence, cost.

The Multiple-Jobs feature may also be used for scoring and analyzing severly entirely different tests, with only a single loading of the program.

REFERENCES

1. Fisher, R.A., *Statistical Methods for Research Workers* (Oliver & Boyd, Edinburgh, 1954), 12th ed., pp. 70–75.

2. Guilford, J.P., *Psychometric Methods* (McGraw–Hill Book Co., Inc., New York, 1954), pp. 373–464.

3. Gullicksen, H., *Theory of Mental Tests* (John Wiley & Sons, Inc., New York, 1950), pp. 363–396.

4. Farrell, R.H., and Stern, G.G., "A Measure of Tetrachoric Association," presented at APA convention, Sept. 1951, Chicago (unpublished; a copy of this paper may be obtained by writing The University of Chicago Computing Center).

5. Wright, B.D., "120X120 Principle Components Analysis: Symmetric (with Varimax Rotation and Plot)," Univ. Chicago Soc. Sci. Div. Computer Library (1963; unpublished).

DESCRIPTIVE TITLE	Zellner's Three-Stage Least-Squares Program
CALLING NAME	Z3SLS NUCC125
INSTALLATION NAME	Vogelback Computing Center Northwestern University
AUTHOR(S) AND AFFILIATION(S)	A. Zellner The University of Wisconsin L. Weiner Department of Industrial Engineering Northwestern University
LANGUAGE	CDC FORTRAN IV
COMPUTER	CDC 6400
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Lorraine Borman, EIN Technical Representative, Vogelback Computing Center, Northwestern University, 2129 Sheridan Road, Evanston, Ill. 60201 Tel.: (312) 492-3682

FUNCTIONAL ABSTRACT

Z3SLS computes two- and three-stage least-squares (2SLS and 3SLS) estimates of parameters in simultaneous or interdependent econometric models. The program first regresses each variable on all the predetermined variables within the model to yield unrestricted least-squares estimates of the reduced-form system. Coefficient estimates, standard errors, etc., are output of the first part of Z3SLS. The second set of computations yields either the 2SLS estimates or both 2SLS and 3SLS estimates, with associated statistics.

REFERENCES

Stroud, A., Zellner, A., and Chou, L.C., "Program for Computing Two- and Three-Stage Least Squares Estimates" (long writeup for Z3SLS), Northwestern Univ. Comp. Ctr. Available from EIN for the cost of duplication and mailing.

DESCRIPTIVE TITLE	FORTRAN Program for Computing a Full Set of Canonical Correlations Relating Two Sets of Measurements
CALLING NAME	CANON
INSTALLATION NAME	University of Pittsburgh
AUTHOR(S) AND AFFILIATION(S)	William W. Cooley, Department of Educational Research, University of Pittsburgh Paul R. Lohnes, Department of Educational Psychology, State University of New York, Buffalo
LANGUAGE	FORTRAN II
COMPUTER	IBM 360/50 IBM 7090
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	**Mr. John Nold, Assistant Director for** Services, Computer Center, University of Pittsburgh, 800 Cathedral of Learning, Pittsburgh, Pa. 15213 Tel.: (412) 621-3500 ext. 7185

FUNCTIONAL ABSTRACT

CANON is a FORTRAN II program that may be utilized for studying the interrelations between two sets of measurements made on the same observational units. Thus, the user may use the routine to test general hypotheses that suggest that relationships exist between two sets of variables. The canonical correlation is the maximum correlation between linear functions of the two sets of variables. Since more than one linear combination of each of the two sets of variables is often possible, the program successively computes pairs of functions subject to the condition that each is orthogonal to all other functions derived from its set of measurements.

The analytical method employed involves the partitioning of the matrix of intercorrelations (R) of the variables into four submatrices.

(1) R_{11} the intercorrelations of the predictor variables

(2) R_{22} the intercorrelations of the criteria variables

(3) R_{12} the intercorrelations of the predictor variables with the criteria variables

(4) the transpose of R_{12}

The four submatrices are then substituted into a canonical equation, the solution of which involves finding latent roots that are functionally related to sets of weights that maximize the correlation between the derived canonical variates.

CANON operates on data cards that contain the full correlation matrix (upper triangular) of order $M = M1 + M2$, where M1 is the number of variables in the left set and M2 is the number of variables in the right set. It is immaterial as to which set represents the predictors or criteria, the only requirement being that $M2 \leq M1$.

CANON reproduces the full correlation matrix that was used as input. In addition, it provides the canonical weights for both predictor and criteria variates, and the factor structures (correlations between each canonical factor and its associated variables) for both sets of canonical factors. Also included are measures of redundancy (the proportion of variance in one set of variables accounted for by the second set of variables) between the two sets of variables, each set given the other.

The Bartlett procedure for testing the significance of canonical correlations using chi square is provided. In addition, χ^2 tests with successive roots removed are provided in the event that the null hypothesis can be rejected.

REFERENCES

Cooley, W. W., and Lohnes, P. R., *Multivariate Procedures for the Behavioral Sciences* (John Wiley & Sons, Inc., New York, 1962; to be republished in 1969).

DESCRIPTIVE TITLE	Computer Means, Median, Frequencies
CALLING NAME	PROFILE NUCC112
INSTALLATION NAME	Vogelback Computing Center, Northwestern University
AUTHOR(S) AND AFFILIATION(S)	Betty Benson, Vogelback Computing Center, Northwestern University
LANGUAGE	CDC FORTRAN IV
COMPUTER	CDC 6400
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Lorraine Borman, EIN Technical Representative, Vogelback Computing Center, Northwestern University, 2129 Sheridan Road, Evanston, Ill. 60201 Tel.: (312) 492-3682

FUNCTIONAL ABSTRACT

PROFILE will give mean, median, sum, and frequency counts for up to 100 *integer-type* variables. Variables may come from 1-, 2-, or 3-column card fields. However, if many of them are 3-column, the number of variables that can be processed will be decreased proportionately.

In tabulating responses, PROFILE discriminates between blanks and punched zeroes. Blanks are considered to be "no response" and consequently are always excluded from the mean and median computation. This is built in and automatic, and it applies to all variables. On the other hand, the user may indicate (using the zero-exclusion card) for each variable whether zeroes are to be excluded or to be treated as "good" answers. For each variable where zeroes are good answers, only a count of *blank* responses will be printed, N (the number of responses) for that variable reduced accordingly, and the sum, mean, and median computed on the basis of the remaining good answers. For each variable where zeroes are excluded, counts of both blank responses and zero responses will be printed, both counts will be subtracted from N for that variable, and the sum, mean, and median will be computed for all *other* answers.

DESCRIPTIVE TITLE	Statistical Analysis of Single Equation Models
CALLING NAME	QSASE
INSTALLATION NAME	Pennsylvania State University Computation Center
AUTHOR(S) AND AFFILIATION(S)	M.C. Hallberg Department of Agricultural Economics and Rural Sociology Pennsylvania State University
LANGUAGE	FORTRAN IV
COMPUTER	IBM 360/67
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Dr. Daniel L. Bernitt, 105 Computer Building, Pennsylvania State University, University Park, Pa. 16802 Tel.: (814) 865-9527

FUNCTIONAL ABSTRACT

QSASE is a self-contained library program designed to compute ordinary least-squares estimates of single equation regression models such as iterative least-squares estimates of special nonlinear regression models (as described in Ref. 1), and two-stage least-squares or limited information maximum likelihood estimates for systems of simultaneous regression equations (as described in Ref. 2). For additional details concerning all options, the user should consult the above references. The following options are available.

1. listing descriptive information about the problem and naming variables
2. using original observations or sums of squares and crossproducts (SSCP) as input data
3. performing various data-editing operations on the input data
4. reading input data from tape or cards
5. printing and/or punching SSCP in either raw form or corrected form
6. printing zero-order correlation matrix

continued

7. printing the mean, standard deviation, standard error of mean, and sum for each variable in the data deck
8. producing a separate listing of all zero-order correlation coefficients that exceed in absolute value a prespecified level
9. producing a scatter diagram of any variable against any other variable
10. stepwise least-squares where a least-squares regression is run and that variable with the smallest t ratio is eliminated in an iterative fashion until all t ratios exceed a prespecified level
11. two-step least squares where a least-squares regression is run and all variables are eliminated whose t ratio do not exceed a prespecified level
12. computing and printing the F ratio for testing the hypothesis that a linear combination of a subset of the $k-1$ β's for a single-equation model are all zero
13. computing and printing the t ratio for testing the hypothesis that a linear combination of a subset of the $k-1$ β's for a single-squation model is equal to a specified value
14. computing and/or printing for each observation the value of the dependent variable Y_t, the estimate of Y_t from the computed regression $\hat{Y}_t$, the estimated error $Y_t-\hat{Y}_t$, and the percentage error
15. producing a scatter diagram of the errors $Y_t-\hat{Y}_t$ against any other variable in the data deck
16. two-stage least-squares estimation
17. limited-information maximum-likelihood estimation

REFERENCES

1. Hallberg, M.C., "Statistical Analysis of Single Equation Stochastic Models Using the Digital Computer," Penn. State Univ. Dept. Agric. Econ. & Rural Sociol. Rep. 78 (February 1969). Copies of this report can be obtained through the contact person.

2. Hallberg, M.C., "Statistical Analysis of System of Simultaneous Linear Equations Using the Digital Computer," Penn. State Univ. Dept. Agric. Econ. & Rural Sociol. Rep. 69 (September 1967). Copies of this report can be obtained through the contact person.

DESCRIPTIVE TITLE	Static Leontief Input—Output Analysis
CALLING NAME	INOUT
INSTALLATION NAME	The Pennsylvania State University Computation Center
AUTHOR(S) AND AFFILIATION(S)	M.C. Hallberg M. Swope Department of Agricultural Economics, The Pennsylvania State University
LANGUAGE	FORTRAN IV
COMPUTER	IBM System 360/67
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Daniel Bernitt, 105 Computer Building, The Pennsylvania State University, University Park, Pa. 16802 Tel.: (814) 865-9527

FUNCTIONAL ABSTRACT

INOUT will solve the static Leontief input—output problem for a model with as many as 175 sectors. From data consisting of a square input matrix, an output vector, and a set of final demand vectors, a great variety of results may be selectively opted, calculated, and printed. Among these are matrices of technical and interdependency coefficients, matrices of interdependency values and net effects, and some other useful vectors.

REFERENCE

Leontief, W., *et al.*, *Studies in the Structure of the American Economy*, Harvard Economic Research Proj. (Oxford University Press, New York, 1953); theoretical and empirical explorations in input—output analysis.

DESCRIPTIVE TITLE	Collection of Statistical Routines
CALLING NAME	STPAC
INSTALLATION NAME	The Pennsylvania State University Computation Center
AUTHOR(S) AND AFFILIATION(S)	William H. Verity Nancy C. Daubert The Pennsylvania State University Computation Center
LANGUAGE	FORTRAN IV
COMPUTER	IBM System 360/67
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	William H. Verity, 104 Computer Building, The Pennsylvania State University, University Park, Pa. 16802 Tel.: (814) 865-9527

FUNCTIONAL ABSTRACT

STPAC consists logically of two sections: a monitor and a collection of statistical routines. The monitor is a program that will read all Control Cards, save the original data if required or requested, save any intermediate results that may be needed later on, and then transfer control to the first routine requested. When this routine is finished, it returns control to the monitor, which then looks to see if there is more work to be done. This work can be of several different types, as follows.

1. Using the same data deck (but possibly different columns of it), execute the same routine again.

2. Using the same data deck, execute another routine. Thus, one deck can be subjected to several types of analyses during one run.

3. Execute another program, possibly the same one as before, on a new set of data.

4. Using the results of a previous routine, continue to another program that takes these results as input; e.g., output from correlation can be the input to regression or factor analysis.

The statistical routines available include a variety for summary

continued

of data, frequency analysis, tests of significance, analyses of variance, and rank and product-moment correlation coefficients. See individual abstracts for EIN Nos. 000 0051(a)-(s) for the details of the individual statistical routines.

REFERENCES

See individual abstracts for EIN numbers 000 0051 (a)-(s).

STPAC ROUTINES

EIN No.	*Calling Name*
000 0051(a)	FAWCS
(b)	STSUM
(c)	TTEST
(d)	PPMCR
(e)	SIGPP
(f)	PARCOR
(g)	CANON
(h)	UPREG/DNREG
(i)	FANAL
(j)	VARMX
(k)	PHICO
(l)	KETAU
(m)	SPRHO
(n)	MNWHT
(o)	KRWAL
(p)	ANOVUM
(q)	ANOVES
(r)	AOVRM
(s)	BARTL

DESCRIPTIVE TITLE	Frequency Analysis with Chi Square
CALLING NAME	(STPAC) FAWCS
INSTALLATION NAME	The Pennsylvania State University Computation Center
AUTHOR(S) AND AFFILIATION(S)	C.B. Broderick Department of Family Relations, The Pennsylvania State University William H. Verity The Pennsylvania State University Computation Center
LANGUAGE	FORTRAN IV
COMPUTER	IBM System 360/67
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	William H. Verity, 104 Computer Building, The Pennsylvania State University, University Park, Pa. 16802 Tel.: (814) 865-9527

FUNCTIONAL ABSTRACT

FAWCS computes (1) discrete, two-dimensional frequency-distribution tables (often called *contingency tables*) whose cells contain an actual frequency and, optionally, (2) expected frequencies and/or percentages. There is no missing-data code inherent to the program, but the user may specify that certain categories of his variables be ignored or combined with one another. If the user has any 2X2 tables, he may request that the Yates correction factor be used in computing his chi-square value.

FAWCS will crosstabulate one variable, the *row* variable, with variables called *column* variables. It can work with up to 50 total variables for any one problem, separately crosstabulating the row variable with each of the column variables. The row variable can have up to 31 categories where 0 is a valid category. Thus, the row variable can be coded 0 through 30. Each column variable can have a maximum of 13 categories where, again, 0 is a valid code. Therefore, column variables can be coded 0 through 12. *Any data outside of these ranges will be ignored.*

continued

REFERENCES

Edwards, A.L., *Experimental Design in Psychological Research* (Holt Rinehart and Winston, New York, 1962).

Snedecor, G.W., *Statistical Methods* (Iowa State University Press, Ames, Ia., 1961).

Wert, J.E., Neidt, C.O., and Ahmann, J.S., *Statistical Methods in Educational and Psychological Research* (Appleton-Century-Crofts, New York, 1954).

DESCRIPTIVE TITLE	Statistical Summary
CALLING NAME	(STPAC) STSUM
INSTALLATION NAME	The Pennsylvania State University Computation Center
AUTHOR(S) AND AFFILIATION(S)	Nancy C. Daubert The Pennsylvania State University Computation Center
LANGUAGE	FORTRAN IV
COMPUTER	IBM System 360/67
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	William H. Verity, 104 Computer Building, The Pennsylvania State University, University Park, Pa. 16802 Tel.: (814) 865-9527

FUNCTIONAL ABSTRACT

STSUM computes, in double-precision arithmetic, the sample size, total, mean, standard deviation, population-variance estimate, standard error, sum of squares, coefficient of variation, third and fourth moments about the origin, third and fourth central moments, alpha (square root of beta 1), alpha 4 (beta 2), momental skewness, and kurtosis. Missing data will be permitted.

Equations

$$\left.\begin{array}{ll} V_1 & = \Sigma X/N \\ V_2 & = \Sigma X^2/N \\ V_3 & = \Sigma X^3/N \\ V_4 & = \Sigma X^4/N \end{array}\right] \quad \begin{array}{ll} \text{moments about origin} & \\ V_1 = \overline{X} & \text{mean} \\ \Sigma X & \text{total} \end{array}$$

$$\left.\begin{array}{ll} \mu_2 & = V_2 - V_1^2 \\ \mu_3 & = V_3 - 3V_1V_2 + 2V_1^3 \\ \mu_4 & = V_4 - 4V_1V_3 + 6V_1^2V_2 - 3V_1^4 \\ \alpha_3 & = \sqrt{\beta_1} = (\mu_3^2/\mu_2^3)^{1/2} \\ \alpha_4 & = \beta_2 = \mu_4/\mu_2 \end{array}\right] \quad \text{moments about mean}$$

$$\begin{array}{rl} \alpha_3/2 & \text{momental skewness} \\ (\alpha_4 - 3)/2 & \text{kurtosis} \\ \Sigma X^2 - (\Sigma X)^2/N & \text{sum of squares} \end{array}$$

continued

REFERENCES

Beyer, W.H., Ed., *Handbook of Tables for Probability and Statistics* (Chemical Rubber Co., Cleveland, 1966), p.3.

DESCRIPTIVE TITLE	T Test on Difference between Means
CALLING NAME	(STPAC) TTEST
INSTALLATION NAME	The Pennsylvania State University Computation Center
AUTHOR(S) AND AFFILIATION(S)	Nancy C. Daubert The Pennsylvania State University Computation Center
LANGUAGE	FORTRAN IV
COMPUTER	IBM System 360/67
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	William H. Verity, 104 Computer Building, The Pennsylvania State University, University Park, Pa. 16802 Tel.: (814) 865-9527

FUNCTIONAL ABSTRACT

TTEST variables (samples) may be correlated or noncorrelated. Missing observations are permitted. Data may be submitted in the form of a single deck or multiple decks. Output can consist of values of t, frequencies, probabilities, and levels of significance. TTEST will perform a t test on each of all possible pairs of variables, each t being computed independently of all others.

For a t test on the difference between means of correlated samples, the formula used is

$$t = \frac{\Sigma D/N}{\left\{[\Sigma D^2 - \frac{(\Sigma D)^2}{N}]/[N(N - 1)]\right\}^{1/2}},$$

where N is the number of pairs and D is the difference between the scores for each pair. The number of degrees of freedom is expressed by N - 1.

For a t test on the difference between means of noncorrelated samples, with separate variances, the formula used is

$$t = (\overline{X}_1 - \overline{X}_2)/(S_{\overline{X}_1}{}^2 + S_{\overline{X}_2}{}^2)^{1/2},$$

continued

where the X's are the means and the $S_{\overline{X}_1}{}^2$ and $S_{\overline{X}_2}{}^2$ are the squares of the standard errors of the means.

If the frequencies of the two samples are equal, the number of degrees of freedom are equal to the frequency minus 1. If the frequencies are unequal, the program uses the lowest frequency minus 1 as the number of degrees of freedom.

A t test on the difference between means of noncorrelated samples, with pooled variances, involves the formula

$$t = (\overline{X}_1 - \overline{X}_2)/[S^2(\frac{1}{K_1} + \frac{1}{K_2})]^{1/2},$$

where S^2 is the variance—i.e., the sum of squares has been calcalculated by summing the squared deviations for each individual case from the mean of the group in which that case is found—and the K's are the individual group frequencies. The number of degrees of freedom is equal to $K_1 + K_2 - 2$.

REFERENCES

Wert, J.E., Neidt, C.O., and Ahmann, J.S., *Statistical Methods in Educational and Psychological Research* (Appleton—Century—Crofts, New York, 1954), pp.128—142.

DESCRIPTIVE TITLE	Pearson Product-Moment Correlation Coefficient
CALLING NAME	(STPAC) PPMCR
INSTALLATION NAME	The Pennsylvania State University Computation Center
AUTHOR(S) AND AFFILIATION(S)	William H. Verity Nancy C. Daubert The Pennsylvania State University Computation Center
LANGUAGE	FORTRAN IV
COMPUTER	IBM System 360/67
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	William H. Verity, 104 Computer Building, The Pennsylvania State University, University Park, Pa. 16802 Tel.: (814) 865-9527

FUNCTIONAL ABSTRACT

PPMCR computes not only a Pearson product-moment correlation coefficient for each possible pair of input variables, but also the mean and standard deviation of each input variable.

DESCRIPTIVE TITLE	Significance of Pearson Product-Moment Correlation Coefficient
CALLING NAME	(STPAC) SIGPP
INSTALLATION NAME	The Pennsylvania State University Computation Center
AUTHOR(S) AND AFFILIATION(S)	W.M. Stiteler, III Mathematics Department, The Pennsylvania State University F. Yates Borden Forestry Department, The Pennsylvania State University H.D. Knoble, The Pennsylvania State University Computation Center
LANGUAGE	FORTRAN IV
COMPUTER	IBM System 360/67
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	William H. Verity, 104 Computer Building, The Pennsylvania State University, University Park, Pa. 16802 Tel.: (814) 865-9527

FUNCTIONAL ABSTRACT

SIGPP is to be used optionally in conjunction with the correlation programs to determine the significance of correlation coefficients at any desired probability levels. That is, the null hypothesis $r = 0$ (alternate: $r \neq 0$) is tested for each correlation coefficient.

REFERENCES

U.S. Department of Commerce, *Handbook of Mathematical Functions* (U.S. Government Printing Office, Washington, D.C., 1964), p.949, formula 26.7.8.

Snedecor, G.W., *Statistical Methods* (Iowa State University Press, Ames, Ia., 1961), p. 190.

DESCRIPTIVE TITLE	Partial Correlation
CALLING NAME	(STPAC) PARCOR
INSTALLATION NAME	The Pennsylvania State University Computation Center
AUTHOR(S) AND AFFILIATION(S)	Gabriele A. Michels The Pennsylvania State University Computation Center
LANGUAGE	FORTRAN IV
COMPUTER	IBM System 360/67
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	William H. Verity, 104 Computer Building, The Pennsylvania State University, University Park, Pa. 16802 Tel.: (814) 865-9527

FUNCTIONAL ABSTRACT

PARCOR constructs the variance–covariance matrix for a given set of variables and then computes the partial correlations and multiple correlations for any specified subset of these variables.

Let A be the variance–covariance matrix for p variables $X_1, X_2, \cdots, X_p$. Let B be the inverse of A. Then, the (p − 2)-order correlation coefficient or the partial correlation coefficient [i.e., the correlation coefficient between X_1 and X_2 holding the other (p − 2) variables fixed] is given by

$$p(12|3\cdots p) = -b_{12}/(b_{11}b_{22})^{1/2}$$

The conditional variance of X_1 is given by

$$\sigma^2(1|23\cdots p) = 1/b_{11}.$$

The multiple correlation coefficient R_1 of X_1 is defined by

$$1 - R_1^2 = \sigma^2(1|23\cdots p)/\sigma_1^2,$$

where σ_1^2 is the variance of X_1. It is estimated by $\hat{R}$ in

$$1 - \hat{R}_1^2 = 1/b_{11}s_1^2,$$

where s_1^2 is the estimated variance of x_1.

continued

000 0051(f)

REFERENCES

Hamilton, W.C., *Statistics in Physical Science* (The Ronald Press Co., New York, 1964), p.184.

000 0051(f)

000 0051(f)

DESCRIPTIVE TITLE	Canonical Correlation
CALLING NAME	(STPAC) CANON
INSTALLATION NAME	The Pennsylvania State University Computation Center
AUTHOR(S) AND AFFILIATION(S)	William H. Verity Nancy C. Daubert The Pennsylvania State University Computation Center
LANGUAGE	FORTRAN IV
COMPUTER	IBM System 360/67
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	William H. Verity, 104 Computer Building, The Pennsylvanir State University, University Park, Pa. 16802 Tel.: (814) 865-9527

FUNCTIONAL ABSTRACT

CANON solves successively for the most significant canonical correlation coefficients. At the same time, two sets of weights associated with each pair of canonical variates are computed.

All matrix computations are done in double-precision floating-point arithmetic. The set of criteria and predictor variables is divided into two groups: l (lefthand) and r (righthand) variables, the smaller set automatically becoming the righthand variables. The program then uses a variation of the jacobian method to extract r eigenvalues and r eigenvectors (righthand weights). These are then used through further calculations to compute r canonical correlation coefficients and r sets of lefthand weights. For each canonical correlation coefficient, chi square is computed with the aid of the eigenvalues and the parameters N (the number of observations), r, and l.

Original variables are brought into the canonical-correlation problem by means of Right- and Left-Variables Cards. Computationally, it is irrelevant whether the variables on the left or on the right are considered as criteria or predictor variables. However, to save computer time, the set with the smallest number of variables (always put into the righthand category by the pro-

continued

gram) should be designated as *right*. Any or all of the variables brought into the problem may be named by using Name Cards according to STPAC rules.

The maximum output records for a problem may be estimated as follows. Let R be the number of righthand variables, L be the number of lefthand variables, and $R < L$. Then, the number of records equals $R(16 + L) + 10$.

Given two sets of variables—not necessarily random variables—[X_i (where $i = 1, 2, \cdots, p$, the number of predictor variables, and Y_i (where $i = 1, 2, \cdots, q$, the number of criterion variables] the problem that CANON solves is essentially that of maximizing the correlation between certain members of the two sets, simultaneously reducing the others to zero. Canonical correlation, then, may be defined in one sense as a process by which the relationship between two sets of variables is reduced to its simplest form. In canonical correlation, *both multiple criteria and multiple predictors* can be involved. (Note that, when $q = 1$ and $p > 1$, the problem is equivalent to that of multiple regression.)

For a hypothetical example, let $p = 3$ and $q = 2$. Here, it is possible to extract two canonical correlation coefficients—say R_1 and R_2. There is a total of $p + q = 5$ variables—say X_1, X_2, X_3 (predictors) and X_4, X_5 (criteria). There will be two pairs of canonical variates—say U_1V_1 and U_2V_2. Suppose for R_1 that the weights for the predictor variables are P_{11}, P_{12}, P_{13} and for the criteria variables that the weights are C_{11}, C_{12}. Then, the pair of new variates, the canonical variates, which are multiples of a linear function of the original variables, are

$$U_1 = K_{11}(P_{11}X_1 + P_{12}X_2 + P_{13}X_3),$$

$$V_1 = K_{12}(C_{11}X_4 + C_{12}X_5),$$

where the K's are arbitrary constants of proportionality and U_1 and V_1 are correlated as described by R_1. A similar situation exists for U_2 and V_2, which are correlated as described by R_2. Furthermore, in general all of the U's and V's have zero mean and unit variance; any U is independent of any other U and likewise for the V's; the correlation between any U and V is zero, except for r correlations $R_1, R_2, \cdots, R_r$, which are the correlations between U_1 and V_1, U_2 and V_2, $\cdots$, U_r and V_r.

As far as interpretation is concerned, the same difficulty arises as in factor analysis: that of knowing whether or not the linear functions correspond to anything real or whether they are merely mathematical figments.

continued

REFERENCES

Cooley, W.W., and Lohnes, P.R., *Multivariate Procedures for the Behavioral Sciences* (John Wiley & Sons, Inc., New York, 1964), pp. 35–59.

Borko, H., *Computer Applications in the Behavioral Sciences* (Prentice-Hall, Inc., Englewood Cliffs, N.J., 1962), pp. 267–279.

DESCRIPTIVE TITLE	Step-Up and Step-Down Multiple Linear Regression
CALLING NAME	(STPAC) UPREG/DNREG
INSTALLATION NAME	The Pennsylvania State University Computation Center
AUTHOR(S) AND AFFILIATION(S)	Dennis Deaven (programming) The Pennsylvania State University Computation Center M.A. Efroymson (flow charts)
LANGUAGE	FORTRAN IV
COMPUTER	IBM System 360/67
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	William H. Verity, 104 Computer Building, The Pennsylvania State University, University Park, Pa. 16802 Tel.: (814) 865-9527

FUNCTIONAL ABSTRACT

UPREG/DNREG computes stepwise multiple-regression equations, adding or deleting one variable at a time until no variables remain to add significantly to the equations or until no significant variables remain in the equation. At each step, the following are printed for those variables currently in the equation.

(1) Regression coefficients
(2) Standard deviation of the regression coefficients
(3) Standard regression coefficients
(4) Partial correlation coefficients
(5) Fraction of explained variance
(6) Multiple correlation coefficient
(7) Regression-equation intercept
(8) Standard deviation of the dependent variable (UPREG only)
(9) F value for the variable entering or leaving the equation (UPREG only)

continued

UPREG

In the step-up procedure, intermediate results are used to give statistical information at each step in the calculation. These intermediate answers are also used to control the method of calculation. A number of intermediate-regression equations are obtained, as well as the final multiple-regression equation. These equations are obtained by adding or deleting one variable at a time, giving the following equations.

$$Y = b\,(0) + b\,(1)x(1),$$
$$Y = b'(0) + b'(1)x(1) + b'(2)x(2),$$
$$Y = b''(0) + b''(1)x(1) + b''(2)x(2) + b''(3)x(3)$$
$$\vdots$$

The variable added is the one that makes the greatest improvement in goodness of fit. The coefficients represent the best values when the equation is fitted by using the specific variables included in the equation. An important property of this procedure is that a variable may be indicated to be significant in any early stage and thus enter the regression equation. In addition, after several other variables are added to the equation, the initial variable may be indicated to be insignificant. The insignificant variable will be removed from the regression equation before adding an additional variable. Therefore, only significant variables are included in the final multiple-regression equation.

DNREG

The step-down procedure of DNREG is simpler than the step-up procedure of UPREG. If the parsimony option has been chosen, DNREG eliminates one variable at a time, printing the intermediate results at each step. The variable eliminated is the one that contributes least to the overall significance. Note that, once a variable has been eliminated with the procedure, it is never brought back into the equation.

REFERENCES

Efroymson, M.A., "Multiple Regression Analysis," in *Mathematical Methods for Digital Computers*, A. Ralston and H.S. Wilf, Eds. (John Wiley & Sons, Inc., New York, 1965).

DESCRIPTIVE TITLE	Factor Analysis (or Principal-Components Analysis)
CALLING NAME	(STPAC) FANAL
INSTALLATION NAME	The Pennsylvania State University Computation Center
AUTHOR(S) AND AFFILIATION(S)	J. Cooley D. Laird L. Pryor J. McConnochie The Pennsylvania State University Computation Center
LANGUAGE	FORTRAN IV
COMPUTER	IBM System 360/67
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	William H. Verity, 104 Computer Building, The Pennsylvania State University, University Park, Pa. 16802 Tel.: (814) 865-9527

FUNCTIONAL ABSTRACT

FANAL solves successively for the most dominant factors represented in a (symmetric) correlation matrix. Factors are ranked according to the variance accounted for. The number of factors to be extracted is specified in advance. The program may also be used to find the dominant eigenvalues and eigenvectors of other (symmetric) positive definite matrices.

All computations are done in double-precision floating-point arithmetic. The method used is an iterative one in which a vector converges to the eigenvector. The largest eigenvalue is found first and then a deflation process is used to find the successively smaller eigenvalues. An extrapolation procedure is used to accelerate convergence of the basic iterative method.

DESCRIPTIVE TITLE	Varimax Rotation
CALLING NAME	(STPAC) VARMX
INSTALLATION NAME	The Pennsylvania State University Computation Center
AUTHOR(S) AND AFFILIATION(S)	J. Cooley D. Thompson John McConnochie The Pennsylvania State University Computation Center
LANGUAGE	FORTRAN IV
COMPUTER	IBM System 360/67
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	William H. Verity, 104 Computer Building, The Pennsylvania State University, University Park, Pa. 16802 Tel.: (814) 865-9527

FUNCTIONAL ABSTRACT

VARMX can perform a number of orthogonal rotations on an arbitrary matrix of factor loadings, using the normal varimax criterion. The result is a unique (within tolerance limits) matrix of factor loadings.

All computations are done in double-precision floating-point arithmetic. First, the original factor loadings as produced by FANAL [EIN No. 000 0051(i)] are normalized by dividing the loadings in each row by the square root of the communality (sum of squares) for that row. The rotation is performed with a criterion of .005 and the resulting loadings are denormalized to give the rotated matrix.

REFERENCES

Kaiser, H.F., "Computer Program for Varimax Rotation in Factor Analysis," Educ. Psychol. Meas. 19, 413–420 (1959).

DESCRIPTIVE TITLE	Phi Coefficient
CALLING NAME	(STPAC) PHICO
INSTALLATION NAME	The Pennsylvania State University Computation Center
AUTHOR(S) AND AFFILIATION(S)	Nancy C. Daubert The Pennsylvania State University Computation Center
LANGUAGE	FORTRAN IV
COMPUTER	IBM System 360/67
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	William H. Verity, 104 Computer Building, The Pennsylvania State University, University Park, Pa. 16802 Tel.: (814) 865-9527

FUNCTIONAL ABSTRACT

PHICO processes a maximum of 105 variables and computes phi coefficients for all possible pairs of variables.

Data are dichotomous and are coded 0 and 1 or 1 and 2. The equation used is

$$\text{phi} = (ad-bc)/[(a+b)(c+d)(a+c)(b+d)]^{1/2}.$$

REFERENCES

Wert, J.E., Neidt, C.O., and Ahmann, J.S., *Statistical Methods in Educational and Psychological Research* (Appleton–Century–Crofts, New York, 1954).

DESCRIPTIVE TITLE	Kendall Rand Correlation Coefficients (tau)
CALLING NAME	(STPAC) KETAU
INSTALLATION NAME	The Pennsylvania State University Computation Center
AUTHOR(S) AND AFFILIATION(S)	Nancy C. Daubert The Pennsylvania State University Computation Center
LANGUAGE	FORTRAN IV
COMPUTER	IBM System 360/67
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	William H. Verity, 104 Computer Building, The Pennsylvania State University, University Park, Pa. 16802 Tel.: (814) 865-9527

FUNCTIONAL ABSTRACT

KETAU computes Kendall rank-correlation coefficients (tau), the values of z scores corresponding to them, and the probabilities and levels of significance of the z scores, for all possible pairs of variables. A correction is made for ties. The correlation coefficients will also be punched on cards, at the option of the user.

REFERENCES

Siegel, S., *Nonparametric Statistics for the Behavioral Sciences* (McGraw—Hill Book Co., Inc., New York, 1956), pp.213—222.

DESCRIPTIVE TITLE	Spearman Rank Correlation Coefficients (rho)
CALLING NAME	(STPAC) SPRHO
INSTALLATION NAME	The Pennsylvania State University Computation Center
AUTHOR(S) AND AFFILIATION(S)	Nancy C. Daubert The Pennsylvania State University Computation Center
LANGUAGE	FORTRAN IV
COMPUTER	IBM System 360/67
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	William H. Verity, 104 Computer Building, The Pennsylvania State University, University Park, Pa. 16802 Tel.: (814) 865-9527

FUNCTIONAL ABSTRACT

SPRHO computes Spearman rank-order correlation coefficients (rho) and the values of t corresponding to them for all possible pairs of variables. A correction is made for ties. The correlation coefficients will also be punched on cards at the option of the user.

REFERENCES

Siegel, S., *Nonparametric Statistics for the Behavioral Sciences* (McGraw-Hill Book Co., Inc., New York, 1956), pp. 202-213.

DESCRIPTIVE TITLE	Mann-Whitney U Test
CALLING NAME	(STPAC) MNWHT
INSTALLATION NAME	The Pennsylvania State University Computation Center
AUTHOR(S) AND AFFILIATION(S)	Nancy C. Daubert The Pennsylvania State University Computation Center
LANGUAGE	FORTRAN IV
COMPUTER	IBM System 360/67
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	William H. Verity, 104 Computer Building, The Pennsylvania State University, University Park, Pa. 16802 Tel.: (814) 865-9527

FUNCTIONAL ABSTRACT

MNWHT computes results for the Mann-Whitney U test and gives corresponding z scores, probabilities of the z scores, and levels of significance. U's are calculated for all possible pairs of variables. Observations per variable may be unequal. A correction is made for ties. U's will be punched on cards at the option of the user.

REFERENCES

Siegre, S., *Nonparametric Statistics for the Behavioral Sciences* (McGraw-Hill Book Co., Inc., New York, 1956), pp. 116-127.

DESCRIPTIVE TITLE	Kruskal–Wallis One-Way Analysis of Variance
CALLING NAME	(STPAC) KRWAL
INSTALLATION NAME	The Pennsylvania State University Computation Center
AUTHOR(S) AND AFFILIATION(S)	Nancy C. Daubert The Pennsylvania State University Computation Center
LANGUAGE	FORTRAN IV
COMPUTER	IBM System 360/67
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	William H. Verity, 104 Computer Building, The Pennsylvania State University, University Park, Pa. 16802 Tel.: (814) 865-9527

FUNCTIONAL ABSTRACT

KRWAL does a Kruskal–Wallis one-way analysis of variance. The levels need not have equal numbers of observations in each. A correction is made for ties.

REFERENCES

Siegel, *Nonparametric Statistics for the Behavioral Sciences*, (McGraw–Hill Book Co., Inc., New York, 1956).

DESCRIPTIVE TITLE	Analysis-of-Variance Method of Unweighted Means
CALLING NAME	(STPAC) ANOVUM
INSTALLATION NAME	The Pennsylvania State University Computation Center
AUTHOR(S) AND AFFILIATION(S)	Nancy C. Daubert The Pennsylvania State University Computation Center
LANGUAGE	FORTRAN IV
COMPUTER	IBM System 360/67
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	William H. Verity, 104 Computer Building, The Pennsylvania State University, University Park, Pa. 16802 Tel.: (814) 865-9527

FUNCTIONAL ABSTRACT

ANOVUM is designed to handle a factorial analysis of variance with either equal or unequal numbers of observations in subclasses. The method of unweighted means is based on the assumption that the population from which the sample is drawn has proportional or equal subclass numbers. It can be relied upon only if the subclass numbers are approximately equal and presumably represent a population with equal numbers. The analysis is performed upon the means for each cell, and each sum of squares is multiplied by the harmonic mean of the number of observations per cell. Up to 150 variables (separate analyses) will be permitted with data code = 1. Only 1 variable is acceptable with data code = 0.

A Bartlett's test for homogeneity of variance is performed on the raw data. If the chi square computed in this test is significant at the 0.05 level, the conclusions of the analysis of variance should be interpreted cautiously. If cell frequencies are greater than 1, and if they are equal, a Scheffe's test will be performed on main effects.

REFERENCES

Winer, B.J., *Statistical Principles in Experimental Design* (McGraw-Hill Book Co., Inc., New York, 1962), pp. 95–96, 222–224.

DESCRIPTIVE TITLE	Analysis of Variance with Unequal Subclass Numbers (Method of Expected Subclass Numbers)
CALLING NAME	(STPAC) ANOVES
INSTALLATION NAME	The Pennsylvania State University Computation Center
AUTHOR(S) AND AFFILIATION(S)	Nancy C. Daubert The Pennsylvania State University Computation Center
LANGUAGE	FORTRAN IV
COMPUTER	IBM System 360/67
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	William H. Verity, 104 Computer Building, The Pennsylvania State University, University Park, Pa. 16802 Tel.: (814) 865-9527

FUNCTIONAL ABSTRACT

ANOVES is designed to handle a factorial analysis of variance with unequal numbers of observations in subclasses (cells). A problem with equal subclass numbers is acceptable, although it is not advisable to use ANOVES for a balanced design with equal subclass numbers owing to its relatively long computing time as compared to that of ANOVUM. Up to 150 variables (separate analyses) will be permitted with data code = 1. Only one variable is acceptable with data code = 0.

This method of analyzing data of multiple classifications with unequal subclass numbers is based on the assumption that the population from which the sample is drawn has proportional or equal subclass numbers. Under its fundamental hypotheses, this method affords an estimate of both the main effects and the interactions. However, every subclass must contain at least one observed value.

If the computed chi-square value for disproportionality is significant, a reduction factor will be calculated and all the sums of squares (except the with-in subclasses, which is calculated from the original data) will be reduced accordingly. This is an attempt to remove part of the disproportionality of the data.

continued

The probability associated with the chi-square value computed by using subclass numbers is calculated and the 0.05 probability point is considered critical in determining whether a reduction factor will be computed or not.

The reduction factor is

$$\frac{N}{\sum_{1}^{K} E^2/A}$$

where N is the total number of observations, K is the number of cell subclasses, E is the expected number of observations in a cell, and A is the actual number of observations in the same cell.

A Bartlett's test for homogeneity of variance is performed on the raw data. If the chi square computed in this test is significant, the conclusions of the analysis of variance should be interpreted cautiously.

REFERENCES

Snedecor, G.W., and Cox, G.M., "Disproportionate Subclass Numbers in Tables of Multiple Classification," Ia. Agriculture Sta. Res. Bul. *180*, 233–272 (1935).

Bennett, K.R. (personal correspondence to J. Streeter).

Winer, B.J., *Statistical Principles in Experimental Design* (McGraw-Hill Book Co., Inc., New York, 1962).

DESCRIPTIVE TITLE	Analysis of Variance with Repeated Measures (Proportionality Assumed)
CALLING NAME	(STPAC) AOVRM
INSTALLATION NAME	The Pennsylvania State University Computation Center
AUTHOR(S) AND AFFILIATION(S)	Nancy C. Daubert The Pennsylvania State University Computation Center
LANGUAGE	FORTRAN IV
COMPUTER	IBM System 360/67
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	William H. Verity, 104 Computer Building, The Pennsylvania State University, University Park, Pa. 16802 Tel.: (814) 865-9527

FUNCTIONAL ABSTRACT

AOVRM is designed to handle a factorial analysis of variance with repeated measures and either equal or unequal numbers of observations in subclasses (cells). A problem with no repeated measures is acceptable, although it is not advisable to use AVORM for such a design owing to the relatively long computing time as compared with that of ANOVM or ANOVES.

This method of analyzing data of multiple classifications with unequal subclass numbers is based on the assumption that the population from which the sample is drawn has proportional or equal subclass numbers. This method allows calculation of all main effects and interactions. However, every subclass must contain at least one subject and the data for every subject included must be complete.

REFERENCES

Winer, B.J., *Statistical Principles in Experimental Design* (McGraw-Hill Book Co., Inc., New York, 1962).

DESCRIPTIVE TITLE	Bartlett's Test for Homogeneity of Variance
CALLING NAME	(STPAC) BARTL
INSTALLATION NAME	The Pennsylvania State University Computation Center
AUTHOR(S) AND AFFILIATION(S)	Nancy C. Daubert The Pennsylvania State University Computation Center
LANGUAGE	FORTRAN IV
COMPUTER	IBM System 360/67
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	William H. Verity, 104 Computer Building, The Pennsylvania State University, University Park, Pa. 16802 Tel.: (814) 865-9527

FUNCTIONAL ABSTRACT

BARTL performs a Bartlett's test for homogeneity of variance on up to 6000 subgroups (or cells). Subgroups may be equal or unequal.

REFERENCES

Winer, B.J. *Statistical Principles in Experimental Design* (McGraw-Hill Book Co., Inc., New York, 1962), pp. 95–96.

DESCRIPTIVE TITLE	Bio-Medical Multivariate Statistical Programs
CALLING NAME	BMD
INSTALLATION NAME	University of Notre Dame Computing Center
AUTHOR(S) AND AFFILIATION(S)	School of Medicine University of California, Los Angeles UNIVAC Division of Sperry Rand Corporation College of Business Administration and Computing Center University of Notre Dame
LANGUAGE	FORTRAN IV
COMPUTER	UNIVAC 1107
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	~~Elizabeth~~ Hutcheson, EIN Technical Representative, Computing Center, University of Notre Dame, Notre Dame, Ind. 46556 Tel.: (219) 283-7784

FUNCTIONAL ABSTRACT

The BMD system is a package of computer programs designed to do both basic data processing and the subsequent statistical analysis. The programs have been prepared in an easy-to-use parametric form so that the researcher may adapt them to a wide variety of statistical problems. For further details of the package, see the User Instructions. The BMD is available at the University of Notre Dame in its 1967 edition.

REFERENCES

Dixon, W.J. (Ed.), *BMD: Biomedical Computer Programs*, (Univ. of Calif. Press, Berkeley, 1967, 1970).

Univ. of Notre Dame Computing Center, *BMD for the UNIVAC 1107*, (Rough draft, 1967 ed. of BMD). Available from Univ. of Notre Dame Computing Center, Notre Dame, Ind.

DESCRIPTIVE TITLE	Matrix Decomposition for Points of View Analysis
CALLING NAME	MATDEC
INSTALLATION NAME	Office of Data Analysis Research Educational Testing Service
AUTHOR(S) AND AFFILIATION(S)	J. Ferris Educational Testing Service
LANGUAGE	FORTRAN IV
COMPUTER	IBM 360/65
PROGRAM AVAILABILITY	Deck and listings presently available
CONTACT	Mr. Ernest Anastasio, Off. of Data Anal. Research, Educational Testing Service, Rosedale Road, Princeton, N.J. 08540 Tel.: (609) 921-9000 ext. 2552

FUNCTIONAL ABSTRACT

The purpose of MATDEC is to decompose a rectangular matrix (i.e., a data matrix, X, dimensioned number of variables by number of subjects) into three matrices U, Γ, and W, according to Horst's[1] development (Ref. 1, pp. 364–382), where Γ is a diagonal matrix of eigenvalues and U and W contain the corresponding eigenvectors. It is intended as a first step in Tucker and Messick's approach to an individual differences model for multidimensional scaling.[2]

As is, MATDEC will handle up to 100 variables or subjects, whichever is the lesser dimension of the data matrix. The program uses F4STAT (FORTRAN IV Statistical System) and, in particular, the routine SDGEXT to develop the characteristic roots and vectors of the crossproducts matrix. Further information about F4STAT can be obtained from Mr. Van Hassel, Educational Testing Service, (609) 921-9000, ext. 2557.

REFERENCES

1. Horst, P., *Matrix Algebra for Social Scientists*, (Holt, Rinehart & Winston, New York, 1963).

2. Tucker, L.R., and Messick, S., "An Individual Differences Model for Multidimensional Scaling," Psychometrika, *28*, 333–367 (1963).

DESCRIPTIVE TITLE	Analysis of Change-over Experiments	
CALLING NAME	ZFE-03, ZFE-04	
INSTALLATION NAME	Office of Data Analysis Research Educational Testing Service (ETS)	
AUTHOR(S) AND AFFILIATION(S)	Procedure:	G. Beall formerly of Gillette Razor Co.
	Program:	V. Halfmann formerly of ETS
LANGUAGE	FORTRAN	
COMPUTER	IBM 360/65	
PROGRAM AVAILABILITY	Decks and listings presently available	
CONTACT	Mr. Ernest Anastasio, Off. of Data Anal. Research, Educational Testing Service, Rosedale Road, Princeton, N.J. 08540 Tel.: (609) 921-9000 ext. 2552	

FUNCTIONAL ABSTRACT

The class of designs covered by this program is latin squares or Youden rectangles (incomplete latin squares). These may be repeated fully or in part. The design may be defective, i.e., certain whole rows may be missing, but no allowance has been made for missing cells, i.e., single observations.

The purpose of this program is primarily to analyze the results for effect of treatment with allowance for carry-over of preceding treatment. There is also direct testing of the significance of carry-over. There is included parallel estimation and testing of significance without allowance for carry-over. The program is self-contained and does not require any external subroutines, such as might be presumed to exist in one form or another at computation centers.

The program, as presently stored, allows analysis for designs up to 40 rows (or blocks). This is limited by the dimensions of the data matrices. It could be readily enough changed to 500 or 1000 if such an experiment were involved. The analysis has been contrived so that such change does not increase the size of the matrix involved in equation solving.

continued

Explanation

The basic equations are,

1. $y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_k + \varepsilon_{ijk}$

where y_{ijk} is an observation assumed built of a general level μ, effect of the i^{th} row or individual α_i, the j^{th} period or column β_j, the k^{th} treatment γ_k and extraneous variability ε_{ijk}. This equation obtains for the first column or period when there has been no conditioning period. For the following periods,

2. $y_{ijk\ell} = \mu + \alpha_i + \beta_j + \gamma_k + \delta_\ell + \varepsilon_{ijk\ell}$

where the effect of the ℓ^{th} treatment in the preceding period is δ_ℓ. For a conditioned experiment Equ. (2) obtains in all columns.

The data actually considered are the differences within rows such as,

$$y_{ijk\ell} - y_{i'j'k'\ell'} = \beta_j - \beta_{j'} + \gamma_k - \gamma_{k'} + \delta_\ell - \delta_{\ell'}$$
$$(j' \neq j\ ,\ k' \neq k\ ,\ \ell' \neq \ell)\ .$$

These differences are then set forth in a matrix. Thus for an unconditioned latin square for which the first line is,

Design:	(1)	(2)	(4)	(3)
Result:	4	5	7	6

we may consider the two differences,

$$y_{111} - y_{1221} = \beta_1 - \beta_2 + \gamma_1 - \gamma_2 - \delta_1 + \varepsilon_{111} - \varepsilon_{1221}$$
$$y_{1221} - y_{1342} = \beta_2 - \beta_3 + \gamma_2 - \gamma_4 + \delta_1 - \delta_2 + \varepsilon_{1221} - \varepsilon_{1342}$$

which results in two lines of the matrix as follows:

β_1	β_2	β_3	β_4	γ_1	γ_2	γ_3	γ_4	δ_1	δ_2	δ_3	δ_4	Result
+1	−1			+1	−1			−1				−1
	+1	−1			+1		−1	+1	−1			−2

Least-squares equations are in the same form. For instance, to get the equation associated with β_2, each line is multiplied by its content in the β_2 column and the product accumulated over all columns. For each set of effects (β, γ or δ), the last equation is replaced by a condition equation,

continued

$$\sum_j \hat{\beta}_j = \sum \hat{\gamma}_k = \sum_\ell \hat{\delta}_\ell = 0 \ .$$

The analysis without carry-over (SANS DELTA) is gotten by replacing, temporarily, all equations appropriate to δ by $\hat{\delta}_\ell = 0$.

The analysis without treatment (SANS GAMMA) is obtained by the temporary replacement, $\hat{\gamma}_k = 0$.

The residual variability is gotten in several steps. First the variability residual on β , γ and δ is

$$3. \quad \sum_{ijk} y^2_{ijk} + \sum_{ijk\ell} y^2_{ijk\ell} - \sum_j \hat{\beta}_j \Big(\sum_{ik} y_{ijk} + \sum_{ik\ell} y_{ijk\ell}\Big)$$

$$- \sum_k \hat{\gamma}_k \Big(\sum_{ij} y_{ijk} + \sum_{ij\ell} y_{ijk\ell}\Big) - \sum_\ell \hat{\delta}_\ell \sum_{ijk} y_{ijk\ell} \ .$$

Secondly, an estimate of $\hat{\mu}$ is made by finding from Equ. (1) or (2) the mean of the values

$$y'_{ijk} = y_{ijk} - \hat{\beta}_j - \hat{\gamma}_k$$

$$y'_{ijk\ell} = y_{ijk\ell} - \hat{\beta}_j - \hat{\gamma}_k - \hat{\delta}_\ell \ .$$

Thirdly, estimates of $\hat{\alpha}_i$ are made by finding the mean of the values

$$y''_{ijk} = y_{ijk} - \hat{\mu}$$

$$y''_{ijk\ell} = y'_{ijk\ell} - \hat{\mu} \ .$$

The residual variability as from equation (3) is then further and finally reduced by

$$\hat{\mu} \sum_{ijk\ell} y_{ijk\ell} + \sum_i \alpha_i \sum_{jk\ell} y_{ijk\ell} \ .$$

Residual variability on the effects of $\hat{\mu}$, rows and columns only (SANS DELTA & SANS GAMMA) is gotten by the formula familiar in analysis of variance.

continued

The test of significance for treatments without allowance for carry-over is based on residual variability SANS DELTA less residual variability SANS DELTA & SANS GAMMA. The test with allowance for carry-over is from residual variability on FULL MATRIX less that on SANS GAMMA. The test for carry-over is from residual variability on FULL MATRIX less that on SANS DELTA.

It need only be added that there is incorporated a test on whether the situation is underdetermined. The program counts the number of different row patterns, multiples this number by the number of columns, and checks whether the result exceeds the number of independent parameters to be estimated. In the case of underdetermination, it refuses to analyze. A second type of refusal arises if the simultaneous equations prove insoluble, which may arise if the design is redundant. Finally, if there is no residual freedom, the program will estimate the parameters but declare F = 0.

DESCRIPTIVE TITLE	Multivariate-Analysis of Variance
CALLING NAME	MANOVA
INSTALLATION NAME	Office of Data Analysis Research Educational Testing Service (ETS)
AUTHOR(S) AND AFFILIATION(S)	Elliot M. Cramer, ETS Amended by: Charles C. Hall, ETS
LANGUAGE	FORTRAN II
COMPUTER	IBM 360/65
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Mr. Ernest Anastasio, Off. of Data Anal. Research, Educational Testing Service, Rosedale Road, Princeton, N.J. 08540 Tel.: (609) 921-9000 ext. 2552

FUNCTIONAL ABSTRACT

MANOVA performs univariate and multivariate analyses of variance, of covariance, and of regression. It is quite general and, initially at least, runs should be coordinated with the Office of Computation Sciences through the contact person until the user becomes familiar with the preparation of parameter cards involved. It will perform univariate and multivariate analyses of variance with complete factorial designs or incomplete designs, and with or without covariates, discriminant analyses and canonical correlations. An exact solution in either the orthogonal or nonorthogonal case is provided, and options include single- or multiple-degree-of-freedom contrasts in the main effect or interactions, transformation of variables, and orthogonal polynomial contrasts with equally- or unequally-spaced points. Reanalyses may be done with different criteria, covariates, contrasts and models.

Limitations include a total of 40 variables, no more than 100 nonvacant cells, and up to eight factors with a maximum of 20 levels in each. There is no limitation on the number of subjects.

continued

REFERENCES

Bock, R.D., "Programming Univariate and Multivariate Analysis of Variance," Technometrics, 5, 95–117 (1963).

Clyde, D.J., Cramer, E.M., & Sherin, R., *Revised Manova Program* (The University of Miami Biometric Laboratory, Coral Gables, Florida, 1967).

Graybill, F.A., *An Introduction to Linear Statistical Models* (McGraw–Hill Book Co., Inc., New York, 1961) 223–253.

Morrison, D.F., *Multivariate Statistical Methods* (McGraw–Hill Book Co., Inc., New York, 1967) 159–206.

Rao, C.R., *Advanced Statistical Methods in Biometric Research* (John Wiley & Sons, Inc., New York, 1952) 236–272.

Roy, S.N., *Some Aspects of Multivariate Analysis* (John Wiley & Sons, Inc., New York, 1957) 76–86.

DESCRIPTIVE TITLE	Factor Analysis by Direct "Oblimin" Method
CALLING NAME	OBLIMIN
INSTALLATION NAME	Educational Testing Service (ETS) Office of Data Analysis Research
AUTHOR(S) AND AFFILIATION(S)	Procedure due to: R. Jennrich, UCLA Program due to: R. Jennrich and H. Harman, ETS Adaptation due to: J. Barone, ETS
LANGUAGE	FORTRAN IV
COMPUTER	IBM 360/65
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Mr. Ernest Anastasio, Office of Data Anal. Research, Educational Testing Service, Rosedale Road, Princeton, N.J. 08540 Tel.: (609) 921-9000 ext. 2552

FUNCTIONAL ABSTRACT

Factor analysis provides that the final solution be in terms of either uncorrelated factors or correlated factors. Beginning in the mid-1940's, following the leadership of Thurstone, there was a trend toward the acceptance of oblique factors. This trend has continued to the present day but, unfortunately, efficient objective means for getting oblique "simple structure" solutions have not generally been available even with modern computers. In 1958, John B. Carroll introduced a whole class of methods for oblique transformation to simple structure. These have come to be known by the term "oblimin" (Ref. 1, pp. 324–326), since they involve *oblique* factors and the *minimization* of a function. The oblimin criterion, which is to be minimized, is given in normalized form by,

$$1. \qquad B = \sum_{p<q=1}^{m} \left[n \sum_{j=1}^{n} (v_{jp}^2/h_j^2)(v_{jq}^2/h_j^2) - \gamma \sum_{j=1}^{n} v_{jp}^2/h_j^2 \sum_{j=1}^{n} v_{jq}^2/h_j^2 \right]$$

continued

The v's in these expressions are the elements of the reference-factor structure matrix, i.e., the correlations between the original variables and the reference factors. Actually, what is desired is to have the primary factor pattern exhibit the principles of simple structure, i.e., large values and near-zero values. In the Thurstone school the "reference structure V" is sought that exhibits the simple structure principles, and then the primary factor pattern is obtained by multiplying the matrix V by a diagonal matrix.

The foregoing indirect, and somewhat awkward, procedure has recently been replaced by a direct approach (Ref. 2). Instead of working with the reference factors that are biorthogonal to the primary factors, Jennrich and Sampson set up a criterion for the direct determination of the primary factors that exhibit the simple structure principles. That criterion may be put in the form,

$$2. \qquad F(A) = \sum_{p<q=1}^{m} \left[\sum_{j=1}^{n} a_{jp}^2 a_{jq}^2 - \frac{\delta}{n} \sum_{j=1}^{n} a_{jp}^2 \sum_{j=1}^{n} a_{jq}^2 \right] ,$$

where A is the matrix of primary factor coefficients. Of course, the loadings may be normalized by rows just as in equation 1. An important difference is that the γ in the indirect method ranges between zero and one, while the δ in equation 2 should be zero or negative.

The object of OBLIMIN is to minimize equation 2. The criterion employed is,

$$\frac{F_{i-1} - F_i}{F_0} \leq \varepsilon ,$$

where i is the iteration number. The output is an oblique factor solution satisfying the principles of simple structure, more or less. When δ is equal to zero, the factors are most oblique. For negative values of δ, the factors become less oblique as δ gets smaller. The solution consists of the factor pattern, the correlations among the factors, and the factor structure. At the present time, the program is limited to n = 100 variables and m = 15 factors.

REFERENCES

Harman, H.H., *Modern Factor Analysis*, Second edition, revised, (University of Chicago Press, Chicago, Ill., 1967)

Jennrich, R.I., & Sampson, P.F., "Rotation for Simple Loadings," *Psychometrika*, 1966, *31*, 313–323.

DESCRIPTIVE TITLE	Unrestricted Maximum Likelihood Factor Analysis
CALLING NAME	UMLFA
INSTALLATION NAME	Office of Data Analysis Research Educational Testing Service
AUTHOR(S) AND AFFILIATION(S)	K. G. Jöreskog Educational Testing Service G. Gruvaeus Educational Testing Service
LANGUAGE	FORTRAN IV
COMPUTER	IBM System 360/65
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Mr. Ernest Anastasio, Office of Data Anal. Research, Educational Testing Service, Rosedale Road, Princeton, New Jersey 08540 Tel.: (609) 921-9000 ext. 2552

FUNCTIONAL ABSTRACT

UMLFA performs a factor analysis of a given correlation matrix. "The factor loadings and the unique variances are estimated by Lawley's method of maximum likelihood. The computational procedure...makes use of the method of Fletcher and Powell for numerical minimization of a function. Any number of factors can be extracted, and each factor is rotated, using Kaiser's varimax method. The goodness of fit of the maximum likelihood solution is tested by Lawley's chi-square test based on the likelihood ratio technique."[1]

The program is able to handle up to 75 variables and 30 factors. However, the input correlation matrix must be positive definite.

REFERENCES

1. Jöreskog, K.G., "A Computer Program for Unrestricted Maximum Likelihood Factor Analysis," Educ. Testing Service Res. Mem. RM-66-20 (1966).

DESCRIPTIVE TITLE	Rao Constellation and Distance Analysis
CALLING NAME	DISCRIM2
INSTALLATION NAME	Michigan State University Computer Institute for Social Science Research
AUTHOR(S) AND AFFILIATION(S)	Stuart Thomas Computer Institute for Social Science Research Michigan State University
LANGUAGE	FORTRAN
COMPUTER	CDC 3600
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Dr. Anders Johanson, Programming Supervisor, Applications Programming, Computer Laboratory, Computer Center, Michigan State University, East Lansing, Mich. 48823 Tel.: (517) 355-4684

FUNCTIONAL ABSTRACT

This program implements the method titled "Constellation and Distance Analysis" when first published by C.R. Rao[1] and later called "Multiple Discriminant Analysis" when presented independently by Bryan[2] and by Lubin. The method operates on a set of variates measured on individuals in several groups. It determines linear combinations of the variates, called discriminant functions, which maximize the ratio of between-group variability to pooled, within-group variability, producing the output listed below. The user can have the program handle data input or write his own subroutine to read data and perform preliminary data manipulations.

Output

1. Job Description
   User comments
   Number of groups and variables
   Options selected

continued

2. Group Summary Statistics

Identification
Number observations specified on group card
Mode of data input and input unit
Format statement (if data read by program rather than user)
First observation in the group (raw and transformed)
Variable means and variances
Variable intercorrelation matrix (optional)

3. Overall Data Summary

Overall means
Overall variances
Overall correlation matrix

4. Discriminant Function Information

Discriminant criterion
Percent trace accounted for by criterion
Rao's chi-square and degrees of freedom for the function
Function weights for raw data
Function weights for data adjusted to unit variances (optional)
Mean discriminant score for each group

5. Overall Discriminant Statistics

Group centroids in discriminant space (optional)
Intercentroid distance matrix (optional)
Back solution of discriminant equation (optional)

6. Discriminant Scores (optional)

Capacity

Number of variables must not exceed 30
Number of groups must not exceed 50
Number of observations in any given group must not exceed 9,999
Number of scores (if requested) must not exceed 10 per respondent

REFERENCES

1. Rao, C.R., *Advanced Statistical Methods in Brometric Research*, (John Wiley & Sons, Inc., New York, 1952) Ch. 9.

2. Bryan, J.G., "The Generalized Discriminant Function: Mathematical Foundation and Computational Routine," Harvard Educational Review, *21*, 2, (1951), 90–95.

DESCRIPTIVE TITLE	Nonmetric Multidimensional Scaling	
CALLING NAME	KRUSKAL	
INSTALLATION NAME	Office of Data Analysis Research Educational Testing Service (ETS)	
AUTHOR(S) AND AFFILIATION(S)	Program due to:	J.B. Kruskal R.F. Long Bell Telephone Laboratories
	Modification due to:	D. Kirk Educational Testing Service
LANGUAGE	FORTRAN	
COMPUTER	IBM 360/65	
PROGRAM AVAILABILITY	Decks and listings presently available	
CONTACT	Mr. Ernest Anastasio, Office of Data Analysis Research, Educational Testing Service, Rosedale Road, Princeton, N.J. 08540 Tel.: (609) 921-9000 ext. 2552	

FUNCTIONAL ABSTRACT

The latest version of the multidimensional scaling program written by J.B. Kruskal of Bell Telephone Labs is available. In addition to the improvements made from earlier versions, a much faster sort has been incorporated into the program and a multi-calculation facility using random starting configurations within one machine run, has been made available. This feature, in addition to being more efficient than the old technique of punching cards for subsequent runs, increases the chances of converging to a global minimum within one machine run with the corresponding savings in time and cost.

A modification to the output was made so that normally only the best of many possible configurations is printed. However, each final configuration and even the individual iterations may be printed (as was done in Kruskal's version) if desired.

The program will handle 60 subjects scaled in up to ten dimensions. Calculations can be repeated up to 99 times on as many as 1800 data values.

continued

REFERENCES

Kruskal, J.B., "Multidimensional Scaling by Optimizing Goodness of Fit to a Nonmetric Hypothesis," Psychometrika, *29*, pp. 1–27, (1964).

Kruskal, J.B., "Nonmetric Multidimensional Scaling: A Numerical Method," Psychometrika, *29*, pp. 115–129, (1964).

DESCRIPTIVE TITLE	Multidimensional Scalogram Analysis
CALLING NAME	MSA-I
INSTALLATION NAME	Office of Data Analysis Research Educational Testing Service
AUTHOR(S) AND AFFILIATION(S)	Procedure due to: L. Guttman J. Lingoes University of Michigan Program due to: J. Lingoes University of Michigan Adaptation due to: D. Kirk Educational Testing Service Advisor on use: H. Harman Educational Testing Service
LANGUAGE	MAP
COMPUTER	IBM 360/65
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Mr. Ernest Anastasio, Off. of Data Anal. Research, Educational Testing Service, Rosedale Road, Princeton, N.J. 08540 Tel.: (609) 921-9000 ext. 2552

FUNCTIONAL ABSTRACT

MSA-I is a program to map types (individuals having the same profile over a set of variables or items) onto an Euclidean space with minimum dimensionality. No assumptions are required about the underlying distributions, the scaling properties of the items, or their ordering. The only requirement is that the categories of each item be mutually exclusive and exhaustive. Types are represented as points in space, each item is a partition of the space, and each category is a region. All types who fall in a given category of a particular item are constrained to be closer to boundary markers of the same category than to delimiters of other categories of the same item. The program we have is a modification of the MSA-I package as distributed

continued

by Dr. Lingoes. An option or modification provides additional plots for each category with the number inserted in the plots corresponding to the response to the item rather than the ID number as in the distributed version.

References and some copies of the substantial literature related to this problem are available through the contact person.

REFERENCE

Lingoes, J.C., "Multiple Scalogram Analysis—A Set-theoretic Model for Analyzing Dichotomous Items," J. of Ed. & Psych. Meas., *23*, pp. 501-524, (1963).

DESCRIPTIVE TITLE	Simulation Package for University Research and Training
CALLING NAME	SPURT
INSTALLATION NAME	Vogelback Computing Center Northwestern University
AUTHOR(S) AND AFFILIATION(S)	Gustave J. Rath Department of Industrial Engineering and Management Sciences Martin Goldberg Leonard Weiner Northwestern University
LANGUAGE	CDC FORTRAN IV
COMPUTER	CDC 6400
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Lorraine Borman, EIN Technical Representative, Vogelback Computing Center, Northwestern University, 2129 Sheridan Road, Evanston, Ill. 60201 Tel.: (312) 492-3682

FUNCTIONAL ABSTRACT

SPURT is a comprehensive package of USASI Standard FORTRAN routines that are designed for use in simulation modelling. These useful routines, ranging from simple to complex, enable the average FORTRAN programmer to employ simulation techniques without having to learn the semantic and syntactic rules of a new programming language.

The SPURT package is made up of six main parts.

I. CLOCK Generation—SPURT1
II. Stochastic Generators—SPURT2
III. Statistical Computations—SPURT3
IV. Analog Simulators—SPURT4
V. List-Processing and Queue-Manipulation—SPURT5
VI. Matrix and Graphical Output—SPURT6

continued

The following is a listing and brief discussion of the various subroutines contained in each of the six SPURT parts.

CLOCK Generation—SPURT1: to implement discrete-time simulation models; to cause events to occur in the proper time sequence

The CLOCK subroutine consists basically of two lists:

Master Time List—contains events scheduled to happen in the future

Master Time Queue—contains events that could not take place at the time when they were scheduled to and, therefore, have been rescheduled; i.e., they have been blocked and are waiting in a queue.

Events can be stored on either list.

CLOCK recognizes two basic kinds of events:

Exogenous—those that are *external* to the user's routine; these are read from Data Cards by the CLOCK

Endogenous—those that are *internal* to the user's routine; these are generated dynamically and then are maintained by the CLOCK

Stochastic Generators—SPURT2: to generate samples from various probability distributions and to calculate sample values

Subroutine	*Usage*
STOGN1	Permits sampling from a discrete empirical probability distribution defined by the user
STOGN2	Enables the user to approximate a continuous distribution by means of a piecewise linear distribution
UNIFRM	Permits the user to sample real values from a uniform distribution in a defined interval
RANDIN	Provides a uniform distribution of integers in a defined interval
NORMAL	Allows the user to obtain a random sample from a normal distribution with given mean and standard deviation
NEGEXP	Permits the user to obtain a random sample from the negative exponential distribution
POISSN	Provides the user with a random sample from the Poisson distribution

continued

ERLANG	Provides a random sample from the Erlang distribution
DISCRT	Permits sampling from a step function describing a discrete cumulative distribution of integer values
LINEAR	Provides the user with a random sample from a cumulative distribution that is obtained by linear interpolation in a nonequidistant table of real values
DRAW	Provides a boolean value of TRUE or FALSE
RANPER	Generates a uniformly distributed, random permutation of the integers 1, 2, ···, M

Statistical Computations—SPURT3: to calculate statistical parameters and histograms of data arrays

Subroutine	*Usage*
STIX1 STIX2 STIX3	Three interrelated subroutines to accumulate and print out a frequency table and to produce a CalComp plot of a normalized histogram of the table
STIX4	Evaluates the mean, standard deviation, maximum value, and minimum value of an array of real numbers
STIX5	Evaluates the correlation coefficient between two arrays of real numbers
STIX6	Ranks an array of real numbers and produces the median and range of the data within the array
STIX7	Produces a statistical description of the data found in an array, including the sample size, mean, standard deviation, standard error, minimum and maximum values, range, and a printed histogram plot

Analog Simulators—SPURT4: to enable the simulation of analog-computer problems on a digital computer

Subroutine	*Usage*
ANALOG SECND	These two subroutines make it possible to obtain output similar to a hybrid computer

continued

List-Processing and Queue-Manipulation—SPURT5: lists are mXn arrays; entries in lists are mX1 arrays.

Subroutine	*Usage*
ADFIFO	Adds an entry at the bottom of the list; it can be removed only after all the elements presently on the list are gone (builds first-in last-out list)
ADLIFO	Adds an entry at the top of the list; it will be removed before any other entry presently on the list (builds last-in first-out list)
REMOVE	Removes the top (or first) entry from a list
PURGE	Destroys the contents of a list
DISPL	Prints the contents of a list

Additional subroutines in SPURT5 provide the capability to rank lists and to delete or to insert entries into lists.

Matrix and Graphical Output—SPURT6: output is facilitated through printing and graphical output

Subroutine	*Usage*
OUT	Prints out a square matrix with column and row headings
NSOUT	Prints out a nonsquare matrix with column and row headings
GRAPH	Produces two-dimensional graphs of plots, using a CalComp plotter

REFERENCES

1. Rath, G.J., "Description of the Simulation Package for University Research and Teaching," Northwestern Univ. Vogelback Computing Ctr., SPURT, Rev. E (1968; unpublished).

DESCRIPTIVE TITLE	Equipercentile Equating Program
CALLING NAME	SCORMACH
INSTALLATION NAME	Wharton Computational Services University of Pennsylvania
AUTHOR(S) AND AFFILIATION(S)	Daniel Ashler, Director Wharton Computational Services University of Pennsylvania
LANGUAGE	FORTRAN IV
COMPUTER	IBM System 360/75
PROGRAM AVAILABILITY	Proprietary; usage permitted but program deck or listing not available
CONTACT	Daniel Ashler, Wharton Computational Services, University of Pennsylvania, Philadelphia, Pa. 19104 Tel: (215) 594-6422

FUNCTIONAL ABSTRACT

SCORMACH uses the equipercentile equating method to obtain "comparable" scores (having identical means, standard deviations, and distributions) for several forms of a given test or for different tests. A given score on one test is considered comparable to a given score on another if, in the two (not necessarily distinct) groups of examinees, the same proportions attain less than the respective given scores—i.e., if a score on one test has the same percentile rank as its comparable score on the other test.

The principal output is a table for each pair of tests by which, given a regressed equipercentile score on either test, one can determine the corresponding regressed equipercentile score on the other. If desired, the same information may also be had in the form of a graph produced on a CalComp plotter. Other tables and graphs allow comparisons of raw scores with raw scores, raw scores with regressed scores, etc.

The scores that are of interest here are the examinees' "true scores" on each respective test. Although true scores cannot be observed directly, the characteristics of the distribution

continued

of true scores for a given test can be estimated. This estimation, resulting in "regressed" scores, is made by modifying the value of each observed score in accordance with the reliability coefficient of the test.[1] One technique for equating scores is to plot points that correspond to pairs of comparable regressed scores on rectangular coordinates. A smooth curve is drawn through the points; pairs of equated scores are read from this curve. In addition to equating scores, the program predicts scores on one test from raw scores on the other. The predicted score is the regressed score that has the same percentile rank as the predictor raw score on the other test.

The program accepts a set of raw scores and a coefficient of reliability for each of two or more tests. (There is no limit to the number of tests to be equated in a single run.) The scores of the first test (*anchor test*) entered as input are equated with the scores of each subsequent test, one test at a time, by the equipercentile method described above.

The comparability of the scores of different forms (or tests) is specific to the type of groups used in obtaining them. A group that is both representative (of the population of examinees for whom the tests are intended) and large (at least 500 examinees) should be used. Preferably, the same examinees should take all tests to be equated (Ref. 1, p. 758). Although the value of the results is enhanced when all tests are taken by the same examinees and by an equal number of examinees, the program does not require that these conditions be met.

REFERENCES

1. Flanagan, J.C., "Units, Scores, and Norms," in *Educational Measurement*, E.F. Lindquist, Ed. (American Council on Education, Washington D.C., 1951), pp. 752–760.

DESCRIPTIVE TITLE	Multiple Scalogram Analysis
CALLING NAME	MSA
INSTALLATION NAME	Office of Computational Sciences Educational Testing Service
AUTHOR(S) AND AFFILIATION(S)	Procedure due to: L. Guttman J. Lingoes University of Michigan Program due to: J. Lingoes University of Michigan Adaptation due to: D. Kirk Educational Testing Service Advisor on use: H. Harman Educational Testing Service
LANGUAGE	MAP
COMPUTER	IBM 360/65
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Mr. Ernest Anastasio, Office of Data Analysis Research, Educational Testing Service, Rosedale Road, Princeton, N.J. 08540 Tel.: (609) 921-9000 ext. 2552

FUNCTIONAL ABSTRACT

This program performs a multiple scalogram analysis using the method of James Lingoes[1]. Do not confuse this with Guttman-Lingoes Multidimensional scalogram analysis which is available under the name of MSA-I (EIN Abstract 000 0070).

REFERENCES

1. Lingoes, J.C., "A Set-Theoretic Model for Analyzing Dichotomous Items," J. of Ed. Psych. Meas., *23*, pp. 501–524, (1963).

DESCRIPTIVE TITLE	Nonmetric Multidimensional Scaling
CALLING NAME	Subroutine KRUSCAL
INSTALLATION NAME	Michigan State University Computer Laboratory
AUTHOR(S) AND AFFILIATION(S)	Alan M. Lesgold Computer Institute for Social Science Research Michigan State University
LANGUAGE	FORTRAN or COMPASS
COMPUTER	CDC 3600
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Dr. Anders Johanson, Programming Supervisor Applications Programming, Computer Laboratory, Computer Center, Michigan State University, East Lansing, Mich. 48823 Tel.: (517) 355-4684

FUNCTIONAL ABSTRACT

KRUSCAL is an implementation of J.B. Kruskal's recently published[1] numerical method for multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis[2]. Given a matrix of similarities or dissimilarities between *n* variables, the routine outputs a configuration of *n* points in a specified number of dimensions such that the distance between any two points is a monotone function of the dissimilarity of the two variables corresponding to those two points.

General Description of Input

KRUSCAL will accept as input any matrix or halfmatrix (below major diagonal) with or without the major diagonal, of similarities or dissimilarities, including, among others, correlation coefficients, confusion probabilities, interaction rates among groups, etc. This matrix need not be symmetric, and the program allows for missing data; interpoint distances corresponding to missing data values do not contribute to the stress.

continued

General Description of Output

The output consists of three sections for the two-dimensional case and two sections for other dimensionalities. In every case, a history of computation is printed, showing for each iteration[1] the following information.

STRESS = Normalized stress achieved[2]

SRAT = The rate of stress improvement = $\frac{\text{STRESS}_i}{\text{STRESS}_{i-1}}$

SRATAV = Weighted average of SRAT

= $\text{SRAT}_i^{1/3} \times \text{SRATAV}_{i-1}^{2/3}$

CAGRGL = Cosine of angle between gradient and previous gradient[1]

COSAV = Weighted average of CAGRGL

= $(1/3)\text{CAGRGL}_i + (2/3)\text{COSAV}_{i-1}$

ACSAV = Weighted average of the magnitude of CAGRGL, computed in the same manner as COSAV

SFGR = Scale factor of gradient, [same as Kruskal's mag(g)][1]

STEP = Step size[1]

Also, in every case, the final number of dimensions is printed. There is an option for printing the interpoint distances in each final configuration, and an option for punching the final configuration onto cards (see description of option cards).

In the two-dimensional case, the final configuration is also plotted (on the line printer, not on the plotter).

REFERENCES

1. Kruskal, J., "Nonmetric Multidimensional Scaling: A Numerical Method", Psychometrika, *29*, pp. 115–129, (June, 1964).

2. Kruskal, J., "Multidimensional Scaling by Optimizing Goodness of Fit to a Nonmetric Hypothesis," Psychometrika, *29*, pp. 1–28, (March, 1964).

DESCRIPTIVE TITLE	Algorithms for Analysis of Variance and Covariance of Incomplete Block and Lattice Designs
CALLING NAME	GAVIAL
INSTALLATION NAME	Iowa State University Computation Center
AUTHOR(S) AND AFFILIATION(S)	K.E. Merritt Statistical Laboratory Iowa State University
LANGUAGE	FORTRAN IV
COMPUTER	IBM 360/65
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Dr. William J. Kennedy, Head, Statistical Numerical Analysis and Data Processing Section, Statistical Laboratory, Iowa State University, Ames, Iowa 50010 Tel.: (515) 294-2260

FUNCTIONAL ABSTRACT

GAVIAL calculates analysis of variance and covariance for a wide variety of statistical problems occurring in the class of incomplete block designs. These include the classes of lattice design.

An algorithm, first proposed by Shah[1], is utilized to solve indirectly the set of reduced normal equations of treatment effects denoted by $A\hat{\tau} = Q$, where the matrix A is $(v \times v)$ of rank $v-1$, and $v \leq 200$ is the number of treatments in the design. Special properties of partially balanced incomplete block designs are exploited to obtain least squares estimates of treatment effects by solving a smaller set of equations denoted by $DZ = L$, where D is $(m \times m)$, and m is the number of distinct associate classes in the design. This method generally results in greater accuracy of estimation and increases the capability for solving larger problems.

GAVIAL utilizes a scan subroutine that simplifies, and for the most part minimizes, the effort required to specify problems to be analyzed. No particular statistical knowledge is required

continued

to use GAVIAL. Generally, a user only needs to know how many treatments or entries, plots, blocks, etc. are present in his experimental design.

REFERENCES

1. Shah, B.V., "A Generalization of Partially Balanced Incomplete Block Designs," Annals of Math. Stat., *30*, pp. 1041–1050, (1959).

DESCRIPTIVE TITLE	An Algorithm for the Optimization of a Quadratic Form Subject to Linear Restraints
CALLING NAME	ZORILLA
INSTALLATION NAME	Iowa State University Computation Center
AUTHOR(S) AND AFFILIATION(S)	D.J. Soults J.J. Zrubeck V.A. Sposito Statistical Laboratory Iowa State University
LANGUAGE	FORTRAN IV
COMPUTER	IBM 360/65
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Dr. William J. Kennedy, Head, Statistical Numerical Analysis and Data Processing Section, Statistical Laboratory, Iowa State University, Ames, Iowa 50010 Tel.: (515) 294-2260

FUNCTIONAL ABSTRACT

ZORILLA will solve quadratic programming problems on the IBM 360 system. The program is composed of a number of subprograms; each is called by a procedure control card. The sequence of control cards defines the solution procedure.

A manual[1] is available to inform the user about the correct formulation for optimizing a quadratic form subject to linear restrictions, and to provide the user with a detailed explanation of how to use the program (order of data deck, key punching format, control cards, etc.).

The program can be used to minimize numerical problems. Keyword commands such as SCAN or MODEL serve as an aid in finding invalid input data or an incorrectly specified model. The program can scale poorly defined problems upon the use of the SCALE agendum card.

continued

ZORILLA will find the values for $x_1, x_2, \ldots, x_n$ which optimize

$$f(x) = p'x - \tfrac{1}{2}\, x'Cx$$

subject to the conditions that

$$Ax \leq b$$
$$x \geq 0$$

where p, C, A contain specified constants. The function f(x) will be referred to as the function which we are optimizing. C must be an n X n symmetric matrix. [Any quadratic form x'Bx may be expressed by a symmetric matrix $x'(\frac{B+B'}{2})x = \frac{1}{2}\, x'Cx$ where $C = (B+B')$.] It is also necessary for C to be a positive (or negative) semi-definite matrix when the objective function f(x) is required to be concave (or convex) in the maximization (or minimization) case.

The program requires a simplex tableau input in the form

$$\begin{array}{cc} & \begin{array}{cc} n & 1 \end{array} \\ \begin{array}{c} n \\ m \end{array} & \left[\begin{array}{cc} -C & -p \\ A & b \end{array}\right] \end{array} \qquad \begin{array}{l} n = \text{number of variables} \\ m = \text{number of restrictions.} \end{array}$$

The identity matrix need not be entered explicitly into the matrix. It is generated by the program.

The simplex method is utilized for transformation. The procedure for choosing the incoming and outgoing vectors is the procedure developed by Van de Panne and Whinston[2]. The method of transforming and updating vectors is the product form of the inverse based on the revised simplex method.

REFERENCES

1. Soults, D.J., Zrubeck, J.J., and Sposito, V.A., *Stat. Num. Anal. and Data Process. Series, No. 9: ZORILLA Reference Manual*, (Statistical Laboratory, Iowa State University, Ames, Iowa, 1969). Available from the Statistical Laboratory at a cost of $0.75.

2. Van de Panne, C., and Whinston, A., "Simplicial Methods for Quadratic Programming," Naval Research Logistics Quarterly, *11*, pp. 273–302, (1964).

DESCRIPTIVE TITLE	A Regression Model Building System
CALLING NAME	MOUFLON
INSTALLATION NAME	Iowa State University Computation Center
AUTHOR(S) AND AFFILIATION(S)	Bonnie L. Hanson Statistical Laboratory Iowa State University
LANGUAGE	FORTRAN IV
COMPUTER	IBM 360/65
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Dr. William J. Kennedy, Head, Statistical Numerical Analysis and Data Processing Section, Statistical Laboratory, Iowa State University, Ames, Iowa 50010 Tel.: (515) 294-2260

FUNCTIONAL ABSTRACT

MOUFLON provides fast and economical use of computational methods for model building in multiple linear regression. The basic model used is

$$y_i = \beta_0 x_{i0} + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_k x_{ik} + e_i$$

$$i = 1, 2, \dots, n$$

where x_{ij}'s are fixed, y_i's are observed values, and e_i's are assumed to be independent $N(0, \sigma^2)$. The values $x_{i0} \equiv 1$, $i = 1, \dots, n$, are usually used. If we substitute $\beta_0 = \bar{y} - \beta_1 \bar{x}_1 - \dots - \beta_k \bar{x}_k$ into the original model, the result is

$$y_i - \bar{y} = \beta_1 (x_{i1} - \bar{x}_1) + \dots + \beta_k (x_{ik} - \bar{x}_k) + e_i$$

or

$$y'_i = \beta_1 x'_{i1} + \dots + \beta_k x'_{ik} + e_i ,$$

which is called the reduced model. The usual procedure solves the reduced normal equations associated with the above model and obtains the constant term by means of the relation

continued

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1\bar{x}_1 - \ldots - \hat{\beta}_k\bar{x}_k \quad .$$

It may be helpful to note that the coefficient matrix corresponding to the reduced normal equations is in fact the matrix of corrected sums of squares and cross products. This coefficient matrix will be called X'X in the future. The following notation is used throughout.

k = total number of independent variables available for inclusion in the regression,

p = number of variables currently included in the regression,

RSS_p = residual sum of squares associated with a p-variate regression

RMS = residual mean sqaure,

REGSS = sum of squares due to regression,

REGMS = regression mean square,

$\hat{\beta}_i^2/c_{ii}$ = reduction in REGSS when ith variable is removed from the regression ($\hat{\beta}_i$ is the estimator of β_i and c_{ii} is the ith diagonal element of $(X'X)^{-1}$

Methods

Four methods are available for selecting the optimal regression.

Hocking and Leslie's Use of the C_p Statistic

Hocking and Leslie's procedure[1] is based on earlier work done by C.L. Mallows.[2,3] Mallows suggested that the selection of a 'good' subset of independent variables in a multiple linear regression be based on the standardized total squared error. He developed the C_p statistic as an estimate of this quantity.

$$C_p = \frac{RSS_p}{\hat{\sigma}^2} - (n - 2p)$$

where $\hat{\sigma}^2$ is the residual mean square obtained by fitting the full model and is used as an estimate of σ^2. Mallows showed that models with small bias tend to have C_p's almost equal to p, and he defined these to be 'good' models. Thus, a graph of C_p versus p will show which of the subsets of independent variates are 'good'. However, this method requires the computation of all regressions.

continued

Hocking and Leslie developed an efficient procedure to find the 'best' (in terms of minimum residual sum of squares) subset or subsets of independent variates. Their approach to the problem is concerned with which subset of r variates, $r = k - p$, should be removed from the regression. It is necessary to note that a minimum reduction in the regression sum of squares due to removing a set of r variables implies that the associated p—variate regression has minimum residual sum of squares. Hocking and Leslie state that an equivalent form of the C_p statistic is

$$C_p = \frac{\text{Red}_p}{\hat{\sigma}^2} + (2p - k)$$

where Red_p is the reduction in regression sum of squares due to removing a set of r variables, where $r = k - p$.

The above equations relate to the reduced model

$$y = \beta_1 + \beta_2 x_2 + \ldots + \beta_k x_k + e$$

where β_1 is the constant term. After adjustment for the notation and model used by the program, the C_p statistic is represented as

$$C_p = \frac{\text{Red}_p}{\hat{\sigma}^2} + (2p - k + 1).$$

The first step of Hocking and Leslie's procedure is to compute the k univariate reductions, θ_i (i.e. θ_i = reduction due to removing the ith variate) and to rank them so that $\theta_1 \leq \theta_2 \leq \cdots \leq \theta_k$. At the same time the variables are relabeled according to the order on the θ_i's. Thus, the removal of the first variable (according to the above labeling) leaves the subset of size $k - 1$ with minimum residual sum of squares among the set of possible subsets of size $k - 1$.

Hocking and Leslie's procedure is based on the following property.

> If the reduction in the regression sum of squares due to eliminating any set of variables for which j is the maximum subscript is not greater than θ_{j+1}, then no subset including any variables with subscripts greater than j can result in a smaller reduction.

The sequential procedure for a given $p = 1, \ldots, k - 2$ (which determine $r = 2, \ldots, k - 1$) is as follows.

continued

Step 1 Compute the reduction R due to removing variables 1,2,...,r (according to above labeling). If $R \leq \theta_{r+1}$, procedure terminates and regression consisting of variables $r + 1,...,k$ is taken to be the 'best' p-variate model. If $R > \theta_{r+1}$, go to Step 2.

Step 2 Include variable $r + 1$ in the set of variates which are candidates for removal. Then compute reductions for all subsets of size r which include variate $r + 1$. Find the smallest of all reductions calculated so far (the reduction found in Step 1 is included in the set). If the minimum reduction is not larger than θ_{r+2}, the procedure terminates and the p-variate set corresponding to the minimum reduction is taken to be the 'best' regression of size p. If the minimum reduction exceeds θ_{r+2}, go to Step 3.

Step 3 Add variable $r + 2$ to the set of variates which are candidates for removal. Compute the reductions for all subsets of size r which include variate $r + 2$. Find the smallest of all reductions calculated so far (the reductions found in Steps 1 and 2 are included in the set). If the minimum reduction is not larger than θ_{r+3}, the procedure terminates and the p-variate set corresponding to the minimum reduction is taken to be the 'best' regression of size p. If the minimum reduction exceeds θ_{r+3}, go to Step 4.

⋮

The procedure continues through as many steps as are necessary to find the subset with minimum residual sum of squares. After the 'best' subset of size p has been found the value of p is decreased by 1 (r increased by 1) and the procedure is executed once again to find the 'best' subset of size p - 1. Once the 'best' subset of size p' (specified by user) is found, the procedure terminates.

Forward Selection

This method[4] assumes the user has a predetermined order for his independent variate. The variates with highest degree of 'importance' for inclusion in the model must be placed first in the ordering and followed by those of lesser 'importance'.

Often, the user will have a basic set of r, $1 \leq r \leq k$, independent variates which are always needed to predict values of the dependent variable y, and a set of $k - r$ variates which might be of value in the prediction equation. He thus wishes

continued

to find out which of the set of k - r variates should be used to predict y. The forward selection procedure begins with the basic set of r variates and proceeds to test each successive variate by means of the following sequential procedure.

Step 1   Test $H_1 : \beta_{r+1} = 0$
Acceptance implies procedure terminates and the basic set of r variates is used to predict y. Rejection implies procedure continues to Step 2 and the basic set of r variates plus variable r + 1 are included in the model.

Step 2   Test $H_2 : \beta_{r+2} = 0$
Acceptance implies procedure terminates and variates $x_1, x_2, \ldots, x_{r+1}$ are used to predict y. Rejection implies procedure continues to Step 3 and variates $x_1, x_2, \ldots, x_{r+1}, x_{r+2}$ are used to predict y.

Step 3   Test $H_3 : \beta_{r+3} = 0$

.

.

.

The basic test of $H_i : \beta_{r+i} = 0$ is to reject if $u > F_\alpha$ where

$$u = \frac{\hat{\beta}_i^2 / c_{ii}}{\text{RMS}}$$

F_α is specified by the user.

Sequential Deletion

This procedure assumes as in Forward Selection[4] that the independent variates' order begins with the 'most important' and ends with the 'least important'. Again, the user specifies a basic set r, $1 < r < k$, of independent variables needed to predict values of $\bar{y}$. The sequential deletion procedure begins with the full k-variate model and tests each preceding variate by means of the following sequential procedure.

Step 1   Test $H_0 : \beta_k = 0$
Rejection implies procedure terminates and all k independent variables are used to predict y. Acceptance implies procedure continues to Step 2 and the kth variate is removed from the model.

Step 2 Test $H_1 : \beta_{k-1} = 0$
Rejection implies procedure terminates and $x_1, x_2, \ldots, x_{r+1}$ are used to predict y. Acceptance implies procedure continues to Step 3 and the (k-1)st variate is removed from the model.

continued

Step 3 Test $H_2 : \beta_{k-2} = 0$

$\vdots$

Step k - r Test $H_{k-r-1} : \beta_{r+1} = 0$
Rejection implies procedure terminates and $x_1, x_2, \ldots, x_{r+1}$ are used to predict y. Acceptance implies procedure terminates but x_{r+1} is removed from the model and only $x_1, x_2, \ldots, x_r$ are used to predict y. Again the test of $H_{k-i} : \beta_i = 0$ is to reject if $u > F_\alpha$ where

$$u = \frac{\hat{\beta}_i^2 / c_{ii}}{\text{RMS}} .$$

F_α is specified by the user.

Stepwise Regression

Before we look at the stepwise procedure[5] let us consider a general case. First, let X_1 denote the set of variates $x_1, \ldots, x_p$ which are currently in the model and let X_2 denote the set of variates $x_{p+1}, \ldots, x_k$.

Next, define the sample partial correlation coefficient of the dependent variable y and one of the independent variates from the set X_2, (say x_{p+i}), to be the simple correlation coefficient of y^* and x_{p+i}^*. y^* is the set of residuals resulting from the regression of y on $x_1, x_2, \ldots, x_p$ and x_{p+i}^* is the set of residuals resulting from the regression of x_{p+i} on $x_1, x_2, \ldots, x_p$. Let the sample partial correlation coefficient of y and x_{p+i} be denoted by r_{p+i}.

Now consider the case of transferring one of the variates, x_{p+j}, from set X_2 to set X_1, (i.e. including x_{p+j} in the regression equation). Let RSS_{p+j} denote the residual sum of squares for the regression of y on $x_1, x_2, \ldots, x_p, x_{p+j}$. Note that the value of j can be $1, 2, \ldots, k-p$.

Let x_{p+i} be 'the' variable transferred to the set X_1 where i satisfies

$$RSS_{p+i} \leq RSS_{p+j} \qquad j = 1, 2, \ldots, k-p$$

This is equivalent to

$$r_{p+i}^2 \geq r_{p+j}^2 \qquad j = 1, 2, \ldots, k-p$$

since it can be shown that

$$RSS_{p+j} = RSS_p(1 - r_{p+j}^2) .$$

continued

Thus, the selection of the largest r^2_{p+j} for $j = 1,2,\dots,k\text{-}p$, minimizes RSS_{p+j}.

Now consider the case of transferring one of the independent variables x_j from the set X_1 to the set X_2 (i.e. removing it from the regression). Let RSS_j denote the residual sum of squares for the regression of y on $x_1, x_2, \dots, x_{j-1}, x_{j+1}, \dots, x_p$. It can be shown that

$$RSS_j = RSS_p + \frac{\hat{\beta}_j^2}{c_{jj}} \ .$$

Let x_i be 'the' variable transferred to the set X_2 where i satisfies

$$\frac{\hat{\beta}_i^2}{c_{ii}} \leq \frac{\hat{\beta}_j^2}{c_{jj}} \qquad j = 1,2,\dots,p$$

or

$$(t^2_c)_i \leq (t^2_c)_j \qquad j = 1,2,\dots,p$$

where $(t^2_c)_i = \frac{\hat{\beta}_i^2/c_{ii}}{RMS}$. Thus, the selection of the smallest $(t^2{}_c)_j$ for $j = 1,2,\dots,p$ minimizes RSS_j.

The above considerations give way to the stepwise regression procedure, which consists of two alternating steps and examination of termination criteria after each step. The procedure terminates when any one of the following criteria is encountered.

1. There is no variable to enter and no variable to remove.
2. The procedure dictates that the same variable be entered and removed successively. This can be corrected by changing the F levels if the user so wishes.
3. The total number of steps executed reaches the maximum number of steps specified by the user.

The procedure begins with Step 1 and no variables entered in the model.

continued

Step 1 Enter variable i into the regression if i satisfies

$$r^2_{p+i} \geq r^2_{p+j} \quad j = 1,2,\ldots,k-p \quad \text{and} \quad (t^2_c)_{p+i} = \frac{(n-p-2)r^2_{p+i}}{(1-r^2_{p+i})} > F_{in} ,$$

where F_{in} is the F level to enter a variable and is specified by the user.

The termination criteria are now checked. If any one of the three criteria is satisfied, the program stops computations. If none of the criteria are satisfied, the program continues to Step 2.

Step 2 Remove variable i from the regression if i satisfies

$$(t^2_c)_i \leq (t^2_c)_j \quad j = 1,2,\ldots,p \quad \text{and} \quad (t^2_c)_i < F_{out} ,$$

where F_{out} is the F level to remove a variable and is specified by the user.

The termination criteria are now checked. If any one of the three criteria are satisfied, the program stops. If none of the criteria are satisfied, the program returns to Step 1.

REFERENCES

1. Hocking, R.R., and Leslie, R.N., "Selection of the Best Subset in Regression Analysis," Technometrics, *9*, 1967, pp. 531—540.

2. Mallows, C.L., *Choosing Variables in a Linear Regression: a Graphical Aid.* Paper presented at the Cent. Reg. Meeting of the Inst. of Math. Stat., Manhattan, Kansas, 1964.

3. Mallows, C.L., *Choosing a Subset Regression.* Paper presented at the Joint Stat. Meeting, Los Angeles, Cal., 1966.

4. Larson, H.J., *Sequential Model Building for Prediction in Regression Analysis*, unpublished Ph.D. thesis, (Iowa State University Library, Ames, Iowa, 1960).

5. Hemmerle, W.J., *Statistical Computations on a Digital Computer*, (Blaisdell Publishing Company, a Division of Ginn and Company, Waltham, Mass., 1967).

continued

6. Garside, M.J., "The Best Subset in Multiple Regression Analysis," Appl. Stat. J. of the Roy. Stat. Soc., *14*, Series C, 1965.

7. Hanson, B.L., *Stat. Num. Anal. and Data Process. Series, No. 12: MOUFLON Reference Manual*, (Statistical Laboratory, Iowa State University, Ames, Iowa, 1969). Available from the Statistical Laboratory at a cost of $0.75.

DESCRIPTIVE TITLE	FURNIVAL Regression Screen
CALLING NAME	FURNIVAL
INSTALLATION NAME	Iowa State University Computation Center
AUTHOR(S) AND AFFILIATION(S)	Van D. Nelson, Numerical Analysis and Programming Section, Statistical Laboratory, Ames, Ia. 50010
LANGUAGE	FORTRAN II
COMPUTER	IBM System 360/65
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Dr. Clair Maple, Director, Computation Ctr., Iowa State University, Ames, Ia. 50010 Tel.: (515) 294-3402

FUNCTIONAL ABSTRACT

Investigators often want to select from a large group of independent variables a smaller number to be used as predictors in a regression equation. One possible method is to compute regressions on *all* possible combinations of the variables. However, the total number of possible regressions can become very large. The purpose of this program is to reduce the number of regressions computed to those combinations of variables meeting one or more of four constraints.

Variables may be fixed or forced to appear in every regression, reducing the number of regressions to those combinations of variables containing the fixed variables. For example, consider the five variables 1,2,3,4, and 5, in which 1 and 3 are to be fixed. Then only those combinations using both 1 and 3 are computed; for instance, the combinations (1,3), (1,3,2), and (1,3,2,5) will have regressions computed, but the combinations (2,4,5), (1,2,4), and (3,5) will not, because *both* 1 and 3 are not present. Also, the maximum number of variables appearing in any regression may be limited to any number less than the total number of variables.

In addition, variables may be placed in sets such that, if one variable in a set appears, all variables in that set will appear. In other words, either the entire set of variables or none of the var-

continued

iables in the set is present. For example, let the number of variables be 5, and let variables 1 and 2 be in set A, variable 3 be in set B, and variables 4 and 5 be in set C. Then regressions would be computed for (3), (1,2), (1,2,3), and (1,2,4,5) but not for (1), (1,3), (3,4), or (2,3,4). Essentially, the program treats a set of variables as a single variable, and instead of using combinations of variables it computes regressions on all combinations of sets.

Variables or sets of variables may also be placed in groups such that,if one member of the group is present in a regression, no other member of that group will be present. For example, using the same sets A, B, and C above, regressions would be computed for (1), (2), (1,3), (1,4), and (2,3,5). Regressions would not be computed on (1,2), (1,2,4), or (1,3,4,5). Either no member or only one member of a group can be present. The same conditions would hold if 1,2,3,4, and 5 represented sets instead of variables.

Computed for each regression are (1) the coefficient of colinearity (the determinant of the gross-moments matrix for the independent variables, scaled by dividing by the logical product of the elements in its main diagonal), used as a measure of the degree of associated among the independent variables, including the dummy variables or constants, and (2) the coefficient of determination (R square), computed conventionally from the formula $R^2 = 1-(SSR/SSY)$, where SSR is the sum of squared residuals and SSY is the sum of squares about the mean of the dependent variable.

REFERENCES

Nelson, Van D., *User Manual—FURNIVAL Regression Screen*, Statistical Lab., Ames, Ia. (1968). Available from the EIN office for the cost of duplication and mailing.

DESCRIPTIVE TITLE	Minres Method of Factor Analysis
CALLING NAME	MINRES
INSTALLATION NAME	Office of Data Analysis Research Educational Testing Service (ETS)
AUTHOR(S) AND AFFILIATION(S)	Program due to: H. Harman, ETS W. Jones, National Security Agency Y. Fukuda, TRW Systems, Inc. Adaptation due to: J. Barone, ETS
LANGUAGE	FORTRAN IV
COMPUTER	IBM 360/65
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Mr. Ernest Anastasio, Office of Data Analysis Research, Educational Testing Service, Rosedale Road, Princeton, New Jersey 08540 Tel.: (609) 921-9000 ext. 2552

FUNCTIONAL ABSTRACT

The word "minres" is a contraction of "minimum residuals," and designates a method of factor analysis involving the minimization of off-diagonal residuals of a correlation matrix. Such a method, which has long been sought, has many features that recommend it for initial factorization of a correlation matrix. While the objective of the principal-factor method is to extract maximum variance, the objective of the minres method is to "best" reproduce the observed correlations. The latter objective can be traced to Thurstone, "The object of a factor problem is to account for the tests, or their intercorrelations, in terms of a small number of derived variables, the smallest possible number that is consistent with acceptable residual errors" (Ref. 1, p. 61). This problem has been resolved (Ref. 2, Chap. 9) by minimizing the residual correlations (i.e., the differences between the observed values and those reproduced from the factor-analysis model).

The basic factor-analysis model may be put in the form

1. $z_j = a_{j1}{-}F_1 + a_{j2}{-}F_2 + \cdots + a_{jm}{-}F_m + d_j{-}U_j \quad (j = 1,2,\cdots,n)$

continued

in which the common factor loadings in the matrix $A = (a_{jp})$ are the only parameters to be estimated. However such a solution is obtained, the matrix $R^{\dagger}$ of reproduced correlations with communalities in the principal diagonal is given by (Ref. 2, p. 28),

2. $R^{\dagger} - AA'$,

where uncorrelated factors are assumed without loss of generality. The condition for a least-squares best fit to the off-diagonal correlations may be expressed as follows.

3. $$f(A) = \sum_{k=j+1}^{n} \Sigma n_{j=1}^{n-1} \left[r_{jk} - \sum_{p=1}^{m} a_{jp} a_{kp} \right]^2 = \min \quad .$$

The objective function in (3) is to be minimized under the constraints.

4. $$h_j^2 = \sum_{p=1}^{m} a_{jp}^2 \leqq 1 \qquad (j=1,2,\cdots,n) \; .$$

Thus, the object of the minres method is to minimize the function f(A) for a specified number of factors m by varying the values of the factor loadings. The communalities (restricted to numbers between 0 and 1) are obtained as a byproduct of the method.

The basic mathematical method employed in the computer program involves the Gauss-Seidel process whereby successive displacements are introduced in only one row of A at a time, making the objective function f(A) quadratic. The mathematical technique for minimizing a quadratic-functions subject to side conditions involving inequalities is extremely difficult. However, the inequalities can be removed from the problem at hand, so that it can be made tractable (Ref. 2, p. 192).

The computation of factor loadings is continued until the following criterion is met,

5. $$\max_{j,p} \left| {}_{(i)}a_{jp} - {}_{(i-1)}a_{jp} \right| < \varepsilon \qquad (j = 1,\cdots,n; \;\; p = 1,\cdots,m) \; ,$$

when i is the iteration number and ε is preset; or until some maximum number of iterations is reached.

The test statistic, U_m (Ref. 2, p. 197), for testing the significance of m factors is calculated at the conclusion of the program. This is asymptotically dustributed as X^2 with $\frac{1}{2}[(n - m)^2 + n - m]$ degrees of freedom, where n is the number of variables.

continued

The computer output includes: (1) input parameters and correlation matrix; (2) initial factor matrix (m principal components; alternative is a principal-factor solution dependent on some communality input); (3) minres solution, along with derived communalities; (4) value of the objective function f(A) for this solution; (5) matrix of residual correlations; (6) frequency distribution of the residuals; (7) frequency distribution of the differences in factor loadings between the final iteration and the preceding one; (8) test of significance for the number of common factors; and (9) time for all the preceding calculations.

REFERENCES

Thurstone, L.L., *Multiple-Factor Analysis*, (University of Chicago Press, Chicago, Ill., 1947).

Harman, H.H., *Modern Factor Analysis*, Second Edition, Revised, (University of Chicago, Ill., 1967).

DESCRIPTIVE TITLE	Analysis of Contingency Tables (ACT version 1.00)
CALLING NAME	G9 *MSU ACT
INSTALLATION NAME	Michigan State University Computer Institute for Social Science Research
AUTHOR(S) AND AFFILIATION(S)	A.M. Lesgold F.M. Sim L.C. Widmayer Computer Institute for Social Science Research Michigan State University
LANGUAGE	FORTRAN and COMPASS
COMPUTER	CDC 3600
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Mr. Anders Johanson, Programming Supervisor, Applications Programming, Computer Laboratory, Computer Center, Michigan State University, East Lansing, Mich. 48823 Tel.: (517) 355-4684

FUNCTIONAL ABSTRACT

This program forms bivariate frequency distributions from designated pairs of variables on decks of punched cards or a magnetic tape. Other terms for bivariate frequency distributions include "cross-tabulations", "cross-runs", "two-way breakdowns", "contingency tables", or merely "tables".

The program may also perform any combination of the following operations on designated tables: row and/or column means and standard deviations; percentages of each cell on the associated row, column and/or table totals; theoretical frequencies; cell contributions to table chi-square and degrees of freedom; contingency coefficient; tau; gamma, product-moment correlation coefficient; and Kruskal-Wallis H. Results are printed with appropriate labeling of variables and types of calculations performed.

continued

Gamma and tau will not be computed in any ordinary ACT runs due to the excessive amount of computation time required.

Limitations

A variable field may be as wide as desired, but each field used must be recoded to a maximum of 100 code groups (0,1,...,99). There is no limit on the number of observations (i.e., the number of cards per data deck) which may be used. The maximum number of variables which may be used is 950. These may be on any number of data decks.

000 0094

DESCRIPTIVE TITLE	Univariate and Multivariate Analysis of Variance, Covariance and Regression
CALLING NAME	NYBMUL
INSTALLATION NAME	Washington University Computing Facilities
AUTHOR(S) AND AFFILIATION(S)	Jeremey D. Finn State University of New York at Buffalo
LANGUAGE	FORTRAN IV
COMPUTER	IBM 360/50
PROGRAM AVAILABILITY	Deck and listing presently available
CONTACT	Dr. C.B. Drebes, Mgr., Scientific Data Processing, Computing Facilities, Box 1098, Washington University, St. Louis, Mo. 63130 Tel.: (314) 863-0100 ext. 3141

FUNCTIONAL ABSTRACT

NYBMUL performs an exact least-squares multivariate analysis of variance or covariance for any crossed and/or nested design. The number of observations per cell may be equal, proportional or disproportionate, including missing observations and incomplete designs. The program can also be used to perform regression analysis, canonical correlation and discriminant analysis.

Input may be either the raw data or a variance-covariance matrix together with means and frequencies. A variety of data transformations are provided which include the use of a matrix transformation. Estimation and analysis phases are based on contrasts which may be specified by the user.

REFERENCES

NYBMUL: Univariate and Multivariate Analysis of Variance and Covariance (Buffalo, N.Y.: SUNY at Buffalo, Comput. Center, 1968).

NYBMUL Manual (Buffalo, N.Y.: Comp. Center Press, SUNY at Buffalo, 1969).

000 0094

DESCRIPTIVE TITLE	University of Maryland Test Scoring Program (Version 4)
CALLING NAME	UOM 32
INSTALLATION NAME	University of Maryland Computer Science Center
AUTHOR(S) AND AFFILIATION(S)	Dr. George Green University of Maryland
LANGUAGE	FORTRAN IV
COMPUTER	UNIVAC 1108
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Mr. Sam Steinberg, Program Librarian Computer Science Center, University of Maryland, College Park, Md. 20740 Tel.: (301) 454-4261

FUNCTIONAL ABSTRACT

This system scores objective tests recorded on Digitek Optical Reader Sheets. Using an answer key it scores tests of up to 160 items for any number of students.

Output

1. A listing of students' answers to each question is optional.

2. For each student, the program lists the number of questions answered right, the number wrong, and the number omitted. Optionally included is a score which imposes a penalty for wild guessing. A standardized T score (which has an arithmetic mean of 50 and a standard deviation of 10) is also printed for each student.

3. A frequency distribution (histogram) and the mean and standard deviation of the scores are printed.

4. An item analysis is printed for *each question* on the examination.

   a) P (Difficulty Index) is the proportion of the total group who answer the question correctly. Items having a

continued

difficulty index of .50 provide maximum differentiation among students. Good tests contain items having difficulty indices in the middle range (.25 to .75).

b) D (Discrimination Index) measures the power of a single item to discriminate between the upper and lower halves of the student group, divided as to total scores. A good classroom test should have indices of discrimination of .30 or better.

Also printed are the responses to each item by the high and low scoring groups. The correct answer is indicated.

The three statistics which follow the D index are used for similar purposes and are consistent in meaning with it. All are discrimination indices expressing the relationship between item success and the total score (the criterion).

c) PHI (Phi coefficient) is an index of discriminating power when the criterion variable is a natural dichotomy and must be used as such; i.e., high-low, good-poor, etc. No assumptions are made about the form of the distribution of the group.

d) RPBI (Point biserial correlation coefficient) is an index of discriminating power when the criterion is a continuous variable. No assumptions are made about the form of the distribution of the group.

e) RB (Biserial correlation coefficient) is an index of discriminating power requiring the assumption that one of the normally distributed underlying variables has been forced into a dichotomy.

The choice among these depends partly on the purpose for which the test and item analysis data are to be used, and partly on the convenience with which each statistic serves that purpose. For most classroom examinations, the D statistic is most easily understood; it will identify items with little internal-consistency discriminating power and lead to greater efficiency of measurement when revising a test to contain the more discriminating items. It should be noted, however, that when RB and RPBI are used, the indices are usually not equal. RB tends to be substantially larger than RPBI.

continued

f) An overall measure of test reliability is provided by Kuder Richardson No. 20 and No. 21 statistics. This is an estimate of how close the same set of scores would result if the same set of items were given again. Most test writers settle with reliabilities over .60 for teacher-made tests. An index of .80 and above reveals a highly reliable examination.

Formulae

Phi Coefficient

$$\emptyset = \frac{BC - AD}{[(A + B)\ (C + D)\ (A + C)\ (B + D)]^{0.5}}$$

Point Biserial Correlation

$$r_{pbi} = \frac{\overline{X}p - \overline{X}_t}{S_t}\sqrt{\frac{p}{q}}$$

$\overline{X}$ = the mean of all scores

$\overline{X}p$ = the mean of all scores in the upper half

$\overline{X}_t$ = the mean of all scores in the lower half

p = proportion of individuals in upper half

q = proportion of individuals in lower half

S_t = standard deviation of all scores

Biserial Correlation

$$r_{bi} = \frac{\overline{X}_p - \overline{X}_t}{S_t} \times \frac{P}{y}$$

P = proportion of cases in the upper group X proportion of cases in the lower group; P = pq

y = height of ordinate of unit normal curve at point of division between p and q

Discrimination Index

$$DIS = \frac{X - Z}{N} \times 2$$

continued

X = total correct answers in upper half

Z = total correct answers in lower half

N = sample size

Kuder-Richardson 20

$$K\text{-}R\ (20) = \frac{NQ}{NQ-1}\ \frac{(S_t^2 - PQ)}{S_t^2}$$

Kuder Richardson 21

$$K\text{-}R\ (21) = \frac{NQ}{NQ-1}\ \left[1 - \frac{\overline{X}_t(NQ-\overline{X}_t)}{NQ\ S_t^2}\right]$$

NQ = the number of questions

PQ = p times q summed over all questions

Correction for Guessing

$$\text{Raw Score} = R - \frac{W}{A-1}$$

R = Rights

W = Wrongs

A = Number of options per item (2 to 5)

REFERENCES

Ebel, R.L., *Measuring Educational Achievements*, (Prentice-Hall, Inc., New Jersey, 1965).

Lindquist, E.F., (Ed.), *Educational Measurement*, (American Council on Education, Washington D.C., 1951).

DESCRIPTIVE TITLE	Spearman Rank-Order Correlation
CALLING NAME	UOM 4
INSTALLATION NAME	University of Maryland
AUTHOR(S) AND AFFILIATION(S)	James A. Forbes University of Maryland
LANGUAGE	FORTRAN II
COMPUTER	IBM 7094
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Sam Steinberg, Prog. Lib., Computer Science Center, Univ. of Maryland, College Park, Md. 20740 Tel.: (301) 454-4261

FUNCTIONAL ABSTRACT

This program computes a Spearman rank-order correlation coefficient and a t ratio to determine whether the correlation coefficient is significantly different from zero. Data are read in a two dimensional matrix where there are K variables with N values of each variable. The portion of the correlation matrix below the diagonal and including the diagonal is not computed. Therefore, where there are K variables, (K!/2!(K-2)!) correlation coefficients are computed. For example, if K=5 the correlation coefficients are computed between variables 1-2, 1-3. 1-4, 1-5, 2-3, 2-4, 2-5, 3-4, 3-5, 4-5.

Output from this job includes the following. The two variables being correlated will be called X and Y in this description.

—the identification numbers of the two variables being correlated,
—the correction factor for the total sum of squares of each variable, based on the number of sets of ties in each variable and the number of ties in each set of ties, (SUMTX, SUMTY, TIESX, TIESY),
—the corrected sum of squares for each variable, (SUMXSQ, SUMYSQ),
—sum of the squared differences between the rank value of X and its corresponding rank value for Y, (SUMDSQ),
—spearman rank-order correlation coefficient, (RHO),
—the total number of values of each variable,
—a t ratio,
—degrees of freedom for the t test.

continued

The following output is optionally available to supplement the summary output described above.

—listing of raw X values with their corresponding raw Y values,
—listing of ranked raw X values with their corresponding Y values,
—listing of the ranked rank of X values with their corresponding Y values,
—listing of ranked raw Y values with their corresponding rank value for X,
—listing of ranked rank values for Y with their corresponding rank value for X,
—listing of difference values—the difference between a rank value of X and its corresponding rank value for Y.

Limitations per problem

The number of X values must not exceed 100.
The number of Y values must not exceed 100.
The number of variables to be intercorrelated (K) must not exceed 100.
The number of tied values in any given set of ties must not exceed 5.

000 0097

DESCRIPTIVE TITLE	Weighted Summing Test Scoring Program, (Recoding)
CALLING NAME	UOM 87
INSTALLATION NAME	University of Maryland
AUTHOR(S) AND AFFILIATION(S)	A.F. Norcio University of Maryland
LANGUAGE	FORTRAN
COMPUTER	UNIVAC 1108
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Sam Steinberg, Prog. Lib., Computer Science Center, Univ. of Maryland, College Park, Md. 20740 Tel.: (301) 454-4261

FUNCTIONAL ABSTRACT

This program selects those variables specified by the user, gives the values of those variables the weights specified by the user and adds the weighted values to a sum. Data may be read from cards or tape. The user is cautioned that the program performs operations that are frequently called recoding and not weighted summation. There is, therefore, a potential semantic difficulty.

Limitations

The user may submit as many separate problems as he desires through the program in only one run. Within each problem the program can handle a maximum of 9,999 observations; for each observation the number of variables must not exceed 999. For each observation the program can compile a maximum of 99 separate sums (scores).

The user may specify up to 1,782 variables for selection; and, it should be noted that any variable may be selected more than once; in addition, a maximum of 891 weights may also be specified.

For each sum (score) that the user desires there must be one Score Card. The user must also specify the format of his data on Variable Format Cards.

continued

000 0097

Summary of Limitations

a) Number of Problems = unlimited
b) Maximum number of Observations = 9,999
c) Maximum number of Variables = 999
d) Maximum number of Weights = 891
e) Maximum number of Selections = 1,782
f) Maximum number of Sums = 99
g) Maximum number of Variable Format Cards = 10

DESCRIPTIVE TITLE	Generalized t Test
CALLING NAME	UOM 5
INSTALLATION NAME	University of Maryland Computer Science Center
AUTHOR(S) AND AFFILIATION(S)	J.F. Williams Computer Science Center University of Maryland
LANGUAGE	FORTRAN IV
COMPUTER	IBM 7090/7094 and UNIVAC 1108
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Mr. Sam Steinberg, Program Librarian, Computer Science Center, University of Maryland, College Park, Md. 20740 Tel.: (301) 454-4261

FUNCTIONAL ABSTRACT

This program computes the t statistic for differences between the means of independent (random) groups or dependent (matched) groups. The user may select to input his data in either of two ways; by rows, where each card represents one variable in all its observations, or by columns, where each card represents one observation with only the variables included in that observation being punched.

Output includes the following.

- sum of values (SUM) for a given variable or group,
- sum of squared values (SUMSQ) for a given variable or group,
- mean of values (XMEAN) for a given variable or group,
- variance of values (SAMVAR) for a given variable or group,
- standard deviation of values (SAMSD) for a given variable or group,
- t (TEE) for a given pair of variables or groups—
- sum of cross-products (SUMCP) and product-moment correlation (R) for a given pair of variables or groups *with a dependent t only,*

continued

—F-ratio (FRATIO) for a given pair of variables or groups *with independent t only.*

Limitations

50 variables of groups (M)
250 observations per variable (N)
3 Variable Format Cards

Computational Formulas

Product-moment correlation coefficient

$$r_{ab} = \sqrt{\frac{[((\Sigma\, X_a X_b/N) - (\Sigma X_a \Sigma X_b))/N^2]^2}{s_a^2\ s_b^2}}$$

Dependent t (using difference scores)

$$t_{ab} = \frac{\overline{X}d}{s_d\sqrt{N}}$$

Independent t

$$t_{ab} = \sqrt{\frac{(\overline{X}_a - \overline{X}_b)^2}{[(N_a s_a^2 + N_b s_b^2) / (N_a + N_b - 2)][(N_a + N_b)/(N_a N_b)]}}$$

REFERENCE

McNemar, Q., *Psychological Statistics*, (John Wiley and Sons, Inc., New York, 1962), 3rd ed.

DESCRIPTIVE TITLE	Individual Case Statistics Program
CALLING NAME	SUMSCRDS
INSTALLATION NAME	The University of Iowa University Computer Center
AUTHOR(S) AND AFFILIATION(S)	Louise R. Levine University of Iowa Computer Center
LANGUAGE	FORTRAN IV
COMPUTER	IBM 360/65
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Mrs. Louise R. Levine, Program Librarian, Applications Programming, University Computer Center, The Univ. of Iowa, Iowa City, Iowa 52240 Tel.: (319) 353-5580

FUNCTIONAL ABSTRACT

SUMSCRDS is used to calculate a mean, sum, standard deviation, and the number of non-missing observations in one individual case. It handles each case separately. The case may continue for as many cards as necessary (no. variables ≠ 80). Missing data are represented as a blank or minus zero.

DESCRIPTIVE TITLE	General Multiple Regression Analysis
CALLING NAME	REGAN1
INSTALLATION NAME	The University of Iowa University Computer Center
AUTHOR(S) AND AFFILIATION(S)	Stan Walljasper, IBM 7044 version Converted to IBM 360 by Louise R. Levine The University of Iowa Computer Center
LANGUAGE	FORTRAN IV (G)
COMPUTER	IBM 360/65
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Mrs. Louise R. Levine, Program Librarian, Applications Programming, University Computer Center, The Univ. of Iowa, Iowa City, Iowa 52240 Tel.: (319) 353-5580

FUNCTIONAL ABSTRACT

REGAN1 computes the Pearson Product Moment correlation matrix, mean, standard deviation, and population standard deviation. If desired, the multiple R, standard error of the estimates, F ratio, degrees of freedom, regression coefficients, partial coefficients and F ratio associated with each individual variate, the intercept constant, and residuals also may be calculated. There is *no* allowance for missing data. The basic equations and a simple explanation may be found in Cooley and Lohnes.[1]

REFERENCES

1. Cooley, W.W., and Lohnes, P.R., *Multivariate Procedures for the Behavioral Sciences*, (John Wiley & Sons, Inc., New York, 1962), pp. 31–45.

DESCRIPTIVE TITLE	Analysis of Variance Using Between-Subject Designs
CALLING NAME	SABCA
INSTALLATION NAME	The University of Iowa University Computer Center
AUTHOR(S) AND AFFILIATION(S)	Dr. Charles Spiker Institute of Child Behavior The University of Iowa Converted for 360/65 by Ted Sjoerdsma The University of Iowa Computer Center
LANGUAGE	FORTRAN IV (G)
COMPUTER	IBM 360/65
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Mrs. Louise R. Levine, Program Librarian, University Computer Center, Univ. of Iowa, Iowa City, Iowa 52240 Tel.: (319) 353-5580

FUNCTIONAL ABSTRACT

This analysis-of-variance routine uses between-subject designs for one, two, or three factors as described by Lindquist[1]. It includes the simple randomized, the two-dimensional factorial (AxB), and the three-dimensional factorial (AxBxC). (*Note*: It includes the 7044 programs SIMRAN, ABFT, and ABCFT.) Proportionality must exist between the corresponding cells (treatment sub-groups) from row to row or from column to column of the table for the AxB; for AxBxC there must be proportionality between the cells of at least two oblongs in the same layer and between corresponding oblongs for all layers. The program can handle up to 830 scores per treatment group and a maximum of 25 treatment groups in each of the A, B, and C dimensions. Input data may be on cards, tape or disk. Output includes a summary table as well as sums, means, variances, and standard deviations for each group. For the AxB and the AxBxC designs, the relevant marginal sums and means are also printed.

SABCA is a between-subject design for one (A), two (A and B), or three (A, B, and C) dimensions; that is, a completely randomized

continued

design concerning inter-subject relations using different subjects for each treatment group. The AxB design can be considered as b repetitions of the simple randomized design where b is the number of treatments in the B dimension. Likewise, AxBxC can be regarded as c repetitions of AxB. Thus AxB is treated as AxBxC with one C treatment, and the simple randomized is treated as AxBxC with one B and one C treatment. The program uses exactly the method described by Lindquist[1] (pp. 49-56, 121-125, Chs. 9 and 10).

Restrictions

Proportionality as described above is required. Otherwise, negative sums of squares may result.

The user should be sure that the sum of squared scores, as well as the other sums, is less than 10^{75}. Otherwise, overflow will result causing an error and incorrect output.

The program handles up to 830 scores per treatment group and a maximum of 25 treatments per dimension.

Machine Requirements

The program runs under 360 operating system and can handle card, disk, or tape input. It requires 19K bytes of core storage.

REFERENCES

1. Lindquist, E.F., *Design and Analysis of Experiments in Psychology and Education*, (Houghton Mifflin Company, Cambridge, Mass., 1953).

DESCRIPTIVE TITLE	Numerical Frequency Analysis
CALLING NAME	NUMFREQ
INSTALLATION NAME	The University of Iowa University Computer Center
AUTHOR(S) AND AFFILIATION(S)	Louise R. Levine The University of Iowa Computer Center
LANGUAGE	FORTRAN IV
COMPUTER	IBM 360/65
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Mrs. Louise R. Levine, Program Librarian, University Computer Center, The Univ. of Iowa, Iowa City, Iowa 52240 Tel.: (319) 353-5580

FUNCTIONAL ABSTRACT

NUMFREQ gives a frequency distribution table similar to those found in most statistical books. It tells how many times a value occurred in the variable and gives the percentage of the total, cumulative percentage, and cumulative frequency for each value. In addition, the mean, standard deviation and standard error of the mean are calculated for each variable. Missing data are represented as a blank or minus zero, or by using missing data codes.

DESCRIPTIVE TITLE	Simple Multiple Linear Regression
CALLING NAME	MISREGN
INSTALLATION NAME	The University of Iowa University Computer Center
AUTHOR(S) AND AFFILIATION(S)	Louise R. Levine University Computer Center
LANGUAGE	FORTRAN IV (G)
COMPUTER	IBM 360/65
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Mrs. Louise R. Levine, Program Librarian, University Computer Center, The Univ. of Iowa, Iowa City, Iowa 52240 Tel.: (319) 353-5580

FUNCTIONAL ABSTRACT

MISREGN performs multiple linear regression, producing means, standard deviations and correlation coefficients. The multiple R, standard error of estimates, F ratio, degrees of freedom of regression, Beta weights, regression coefficients and y intercept are calculated from a chosen subset of the correlation matrix. Residuals may be computed. Correlations are calculated using a pairwise technique. Fifty variables are allowed. Missing data are allowed in the form of blanks or minus zero (-0).

Caution: The distortion due to the missing values may cause the results of multivariate analysis to be meaningless.

REFERENCES

Cooley, W.W., and Lohnes, P.R., *Multivariate Procedures for the Behavioral Sciences*, (John Wiley & Sons, Inc., N.Y., 1962), pp. 31-35.

DESCRIPTIVE TITLE	Test-Scoring and Item-Analysis Package
CALLING NAME	(a) QUICKSCORE; (b) ITEMSTEP; (c) ITEMRS
INSTALLATION NAME	Wharton Computational Services Wharton School of Finance and Commerce University of Pennsylvania
AUTHOR(S) AND AFFILIATION(S)	Daniel Ashler Daniel Bricklin David Sheinson Wharton Computational Services
LANGUAGE	FORTRAN
COMPUTER	IBM 360/75
PROGRAM AVAILABILITY	Available to process data at Wharton Computational Services
CONTACT	Daniel Ashler, Wharton Computational Services, Wharton School of Finance and Commerce, Dietrich Hall, University of Pennsylvania, Philadelphia, Pa. 19104 Tel.: (215) 594-6422

FUNCTIONAL ABSTRACT

The Test-Scoring and Item-Analysis Package currently consists of three programs, QUICKSCORE, ITEMSTEP, and ITEMRS. Together, they provide a flexible facility for tasks that range from simple scoring of tests to the most sophisticated test analysis currently available. The best features of item-analysis programs in use on various campuses have been incorporated.

QUICKSCORE

...is the least expensive of the three programs to use. It scores tests and lists the examinees and their scores, first in alphabetical order and then in order of score. Up to 500 examinees can be scored at once. However, the input can be "batched;" that is, several groups of cards of up to 500 each can be scored in this manner, one after another. Beside each examinee's name is printed his Social Security number, his score (which has been corrected

continued

for chance success), the number of items that he got correct, that he got incorrect, that he omitted, and the number of items not reached. The corrected score gives +1 for each item answered correctly and $-1/(c-1)$ for each answered incorrectly, where c is the number of choices per item. Omitted items are scored 0.

A brief item analysis is then performed for each question. First, the estimated fraction of the examinees who knew the right answer is computed (difficulty of question). The examinees are then partitioned into an upper and a lower group on the basis of total test score, these groups being equal or nearly equal in size. A 2X2 table is constructed as follows.

	No. Rights	No. Wrongs
Upper half	A	B
Lower half	C	D

Two indices are computed on this table: ϕ correlation coefficient (PHI) and a ratio index (RI). For a test with N choices per question and U examinees in the upper half and L examinees in the lower, the two indices are defined as follows.

$$\text{PHI} = \frac{AD - BC}{[(A+B)\ (A+C)\ (C+D)\ (B+D)]^{\frac{1}{2}}}$$

$$\text{RI} = \frac{2}{\pi} \arctan \log_2 (\text{UPSC}/\text{LOWSC}),$$

where

$$\text{UPSC} = \frac{A - B/(N-1)}{U} + \frac{1}{N-1}$$

and

$$\text{LOWSC} = \frac{C - D/(N-1)}{L} + \frac{1}{N-1}\,.$$

Both indices can have values in the range −1 to 1. The ratio index reflects the extent of discrimination between the upper and lower groups as well as the correlation of the score on the given question with total test score.

A vertical histogram of the score distribution follows, in which, however, minus scores, scores of zero, and scores of one are all counted in the score category of one. Test statistics follow, in which are given the number of questions in the test, the number of choices per item, the number of examinees, a copy of the scoring key, and the means, standard deviations, and variances of the scores.

continued

ITEMSTEP

...not only performs the same scoring as QUICKSCORE but also provides fuller item analysis and a stepwise-reduction feature. After scoring and analysis, one or more questions (items) are deleted on the basis of some criterion (supplied as a parameter), after which the shortened test is rescored and reanalyzed. The cycle of shortening and reprocessing is repeated as many times as specified. One of three different indices may be specified, on the basis of which the less desirable items shall be eliminated. The first is the W index, which is a weighted combination of the Davis[1] difficulty and discrimination indices. The second is the P-adjusted index, which is the proportion of correct ~~responses,~~ adjusted or corrected for chance success. The third is the contribution of the item to the total test variance. The user specifies either the index level required for retention of an item or that some fixed number of items with the worst index values be dropped at each cycle. A reliability coefficient is also computed. The stepwise-reduction process tends to produce successive score distributions that are more and more rectangular.

ITEMRS

...is the most advanced item-analysis program presently available. First, it provides a header page, in which the indices are explained. Then there are optional printouts of the main matrices produced during the run: namely, the sum-of-products matrix, variance–covariance matrix, and the product–moment correlation matrix. (It should be noted that the correlations are product–moment and not point biserial because the items are not scored dichotomously.) Next, an optional cluster analysis is performed on the correlations after they have been transformed to Fisher's Z's. (A cluster analysis is quicker and less expensive than factor analysis and for many purposes is just about as useful.) Each pass produces larger clusters and fewer of them.

Next, the examinee's scores are listed in a tabulation like that of QUICKSCORE, in order by name and then by score. Alongside the listing of the scores are 16 blank columns. If desired, data may be printed in these columns, such as scaled scores, grades, or comments. (For example, to print scaled scores, one provides a card for each possible score that can be obtained, specifying what shall be printed for that score.)

A large number of indices are computed for each item, including P (proportion of correct responses) and Q (proportion of incorrect responses), both unadjusted and adjusted for chance success; the variance contribution of the item; the product–moment correlation of the item with the total score; the "latent" correlation, which

continued

is the estimate of the correlation between the underlying ability to answer the item correctly and the total score; and the Fisher's Z transform of the product–moment correlation. Also, the effective number of choices (that is, the number of choices that actually attracted examinees) the estimated proportion who knew the answer (based on the effective number of choices), the Davis difficulty and discrimination indices, the W index, which is a weighted combination of the Davis indices, and the index DICAP, which measures the deviation of the responses to this item from the ideal pattern in which 50% of the examinees would choose the correct alternative and the remaining 50% would distribute equally over the remaining alternatives. Optionally, a tabulation of the number of responses to each alternative is given for each fifth of the examinees.

Following the item analysis, a histogram is printed, followed by test statistics, including the covariance reliability coefficient, the KR-20 reliability coefficient, the standard error of measurement, the variance error of measurement, and the test mean, standard deviation, and variance of total test scores.

REFERENCES

1. Davis, F.B., *Item Analysis Data* (Harvard University Press, Cambridge, Mass., 1949).

DESCRIPTIVE TITLE	Smallest Space Analysis
CALLING NAME	SSA-1
INSTALLATION NAME	Office of Data Analysis Research Educational Testing Service
AUTHOR(S) AND AFFILIATION(S)	Procedure due to: L. Guttman J. Lingoes University of Michigan Program due to: J. Lingoes University of Michigan Adaptation due to: D. Kirk Educational Testing Service
LANGUAGE	MAP
COMPUTER	IBM 360/65
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Mr. Ernest Anastasio, Off. of Data Anal. Research, Educational Testing Service Rosedale Road, Princeton, N.J. 08540 Tel.: (609) 921-9000 ext. 2552

FUNCTIONAL ABSTRACT

SSA-1 represents a nonmetric technique for finding the smallest Euclidean space for a configuration of points. To quote from the author's description,

> Briefly stated, the problem posed for the...program is: given a matrix of inequalities among pairs of points in a metric or nonmetric space, determine a set of Euclidean coordinates such that the distances calculated from then are a monotonic function of the ranks or order among the inequalities.

(The voluminous list of references related to this topic are available through the contact person.)

According to the authors, a major computational advantage of this program is that it avoids the local minima problem inherent in the Kruskal program (see EIN No. 000 0068) by employing a rank-image principle. The program can handle up to 70 variables in ten dimensions.

DESCRIPTIVE TITLE	Best Least Squares Polynomial Approximation Subroutine
CALLING NAME	LSPOL
INSTALLATION NAME	The University of Iowa University Computer Center
AUTHOR(S) AND AFFILIATION(S)	James Hurt Department of Mathematics University of Iowa
LANGUAGE	FORTRAN IV
COMPUTER	IBM 360/65
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Mrs. Louise R. Levine, Program Librarian, University Computer Center, The Univ. of Iowa, Iowa City, Iowa 52240 Tel.: (319) 353-5580

FUNCTIONAL ABSTRACT

The LSPOL routine computes the coefficients of the polynomial of degree n which best approximates (in the least squares sense) a set of data. The degree of the polynomial must be less than or equal to 20. The method used is that of Beckett and Hurt[1].

REFERENCES

1. Beckett, R., and Hurt, J., *Numerical Calculations and Algorithms*, (McGraw-Hill Book Co., Inc., New York, 1967), pp. 264-290.

DESCRIPTIVE TITLE	Stepwise Regression Analysis
CALLING NAME	STEPREGN
INSTALLATION NAME	The University of Iowa University Computer Center
AUTHOR(S) AND AFFILIATION(S)	Louise R. Levine The University of Iowa Computer Center Revision of BMD02R program
LANGUAGE	FORTRAN IV
COMPUTER	IBM 360/65
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Louise R. Levine, Program Librarian, University Computer Center, The Univ. of Iowa, Iowa City, Iowa 52240 Tel.: (319) 353-5580

FUNCTIONAL ABSTRACT

STEPREGN computes a sequence of multiple linear regression equations in a stepwise manner. At each step one variable is added to the regression equation. The variable added is the one which makes the greatest reduction in the error sum of squares. In addition, it is the variable which has highest partial correlation with the dependent variable partialed on the variables which have already been added; and it is the variable which, if it were added, would have the highest F value. Variables can also be forced into the regression equation. Non-forced variables are automatically removed when their F values become too low. Regression equations with or without the regression intercept may be selected.

Limitations per problem:

p	number of original variables ($2 \le p \le 80$)
q	number of variables added by transgeneration ($-9 \le q \le 78$)
p+q	total number of variables ($2 < p+q \le 80$)
s	number of Sub-problem Cards ($1 < s < 99$)
k	number of Variable Format Cards ($1 < k < 10$)
i	number of variables to be plotted ($0 \le i \le 30$)
n	number of cases ($1 \le n \le 9999$)
m	number of Transgeneration Cards ($0 \le m \le 99$)

continued

Estimation of output pages per problem.

$$\text{Number of pages} = \frac{\text{no. of steps } [23 + (3/4)(p+q)]}{56} + 5 \text{ per sub-problem.}$$

Transgeneration of the variables is allowed (see Transgeneration Codes).

REFERENCE

Dixon, W.J., (Ed.), *BMD: Biomedical Computer Programs*, (Univ. of Calif. Press, Berkeley, Calif., 1967), pp. 233-259d.

DESCRIPTIVE TITLE	General Program for Multivariate Cross-classification
CALLING NAME	NUCROS
INSTALLATION NAME	The University of Iowa University Computer Center
AUTHOR(S) AND AFFILIATION(S)	K. Janda Northwestern University Adapted for IBM 7044: Merle Wood, Political Science Department, University of Iowa Converted for IBM 360: Janice Lewis, University Computer Center, University of Iowa
LANGUAGE	FORTRAN IV (G)
COMPUTER	IBM 360/65
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Mrs. Louise R. Levine, Program Librarian, University Computer Center, The Univ. of Iowa, Iowa City, Iowa 52240 Tel.: (319) 353-5580

FUNCTIONAL ABSTRACT

NUCROS prepares multivariate cross-classifications in accordance with user-supplied instructions. It will compute row and column sums, and if the user desires, it will also compute row and column percentages, the chi square value, the contingency coefficient, and the Kendall tau-b or tau-c, the Goodman-Kruskal gamma, and Somers D_{yx} and D_{xy} correlation coefficients. There may be up to 40 variables and up to 72 tables processed for each cycle, with a maximum of 99 cycles. Two variables are cross-classified in each table, but a third and fourth variable may be held constant. The data are assumed to be positive integers.

REFERENCES

Janda, K., *Data Processing*, (Univ. of Northwestern Press, Evanston, Ill., 1964).

DESCRIPTIVE TITLE	Data Survey and Normality Test
CALLING NAME	NORMSURV
INSTALLATION NAME	The University of Iowa University Computer Center
AUTHOR(S) AND AFFILIATION(S)	Carol Hopkins Yale University Computer Center Revised by Louise Levine The University of Iowa Computer Center
LANGUAGE	FORTRAN IV
COMPUTER	IBM 360/65
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Louise R. Levine, Program Librarian, University Computer Center, The Univ. of Iowa, Iowa City, Iowa 52240 Tel.: (319) 353-5580

FUNCTIONAL ABSTRACT

NORMSURV allows the user to test the hypothesis that his data are normally distributed, using a chi-square test. Missing data are detected as either minus zero or blank in the input data. Values of zero are considered to be valid and are used in all calculations. Means, standard deviations and a list of outliers (see below) will be indicated on the output. Transformations are allowed to generate new variables or to change any variable before the analysis is done. A median and test of skewness are also calculated.

The data are scanned variable by variable for missing data. Means and standard deviations are calculated using only the number of non-missing observations. The data are then standardized (mean = 0, variance = 1), checked for extremes, and placed in a 10-cell frequencies distribution. The limits of the cells are the deciles of the normal distribution of standardized data. The estimate of chi-square has seven (7) degrees of freedom because three degrees are lost in estimating μ by $\bar{x}$, σ by s, and forcing $\sum_{i=1}^{10} f_e = k_n$ where f_e is the expected frequency in each cell.

continued

Output includes the following:

—list of transformation cards,
—list of data before and after transformation (optional),
—list of missing observations,
—number of non-missing observations, sample mean and standard deviation
—list of observations whose standard scores are greater in absolute value than an extreme value of the user's choice (outliers),
—a graph of the cell frequencies
—an estimated chi-square statistic.

DESCRIPTIVE TITLE	Fitting Nonlinear Regression Functions
CALLING NAME	TARSIER
INSTALLATION NAME	The University of Iowa University Computer Center
AUTHOR(S) AND AFFILIATION(S)	J.D. Atkinson Statistical Laboratory Iowa State University
LANGUAGE	FORTRAN
COMPUTER	IBM 360/65
PROGRAM AVAILABILITY	Available for use; not distribution
CONTACT	Louise R. Levine, Program Librarian Applications Programming, University Computer Center, The University of Iowa, Iowa City, Iowa 52240 Tel.: (319) 353-5580

FUNCTIONAL ABSTRACT

TARSIER is a programmed method for the fitting of nonlinear regression functions. The main program implements Hartley's[1] Modified Gauss—Newton Method for Fitting of Non-Linear Regression Functions by Least Squares. A subroutine, FCNT, is written for each individual application of TARSIER, describing the function to be fitted.

The Modified Gauss—Newton Method (MGN) is iterative and requires a starting value for each parameter to be estimated. The efficiency of the MGN method depends on the reasonable selection of the starting values.

REFERENCES

1. Hartley, H.O., "The Modified Gauss-Newton Method for the Fitting of Non-Linear Regression Functions by Least Squares," Technometrics, Vol. 3, No. 2 (May 1961), pp. 269—280. Reprint Series No. 90, Stat. Lab., Iowa State University.

continued

2. Hartley, H.O., and Booker, A., "Nonlinear Least Squares Estimation," Annals of Math. Stat., 36:2 (April 1965), pp. 638–650. Reprint Series No. 165, Stat. Lab., Iowa State University.

3. Faddev, D.K., *Computational Methods of Linear Algebra* (New York: Dover Publications Inc., 1959) pp. 105-111.

4. Atkinson, J.D., "TARSIER," (Ames, Iowa: Iowa State University Statistical Laboratory, 1969). Available from the EIN Office at the cost of reproduction and mailing.

DESCRIPTIVE TITLE	Student's t Test
CALLING NAME	STUDENTT
INSTALLATION NAME	The University of Iowa University Computer Center
AUTHOR(S) AND AFFILIATION(S)	Modified version version of BMDX70 (T Program) from U.C.L.A. Louise R. Levine University Computer Center The University of Iowa
LANGUAGE	FORTRAN IV
COMPUTER	IBM 360/65
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Mrs. Louise R. Levine, Program Librarian, Applications Programming, University Computer Center, The University of Iowa, Iowa City, Iowa 52240 Tel.: (319) 353-5580

FUNCTIONAL ABSTRACT

This program computes t statistics and associated probability levels to test the equality of the means of two groups based on pooled and separate variance estimates. An F statistic and associated probability level for the equality of group variances is also computed. Paired comparison t ratios may be obtained through transgeneration. Groups are defined in two possible ways: 1) variable vs. variable, or 2) partitioned variable. Several dependent variables may be analyzed simultaneously. Each problem may contain from one to twenty subproblems. The cases to be included in each subproblem are determined through Boolean selection. Transgenerations are available and data specified as "missing" will be deleted. Variables transgenerated from variables with missing values will be considered missing.

Output

Output from this program includes F ratio of variances, t value (based on pooled variance estimate), t value (based on separate variance estimate), two-tailed probability levels for each t

continued

and for the F, means, standard deviations, standard error of the means, and number of observations included in the computation of the means, standard deviations, and standard error of the means.

Limitations per problem

p — number of original variables ($1 \le p \le 199$)

n — number of cases ($1 \le n \le 32{,}000$)

q — number of variables added to the original set after transgeneration ($-198 \le q \le 99$)

p+q — total number of variables output ($1 \le p+q \le 100$)

m — number of Transgeneration Cards ($0 \le m \le 100$)

D — number of Missing Value Cards ($0 \le D \le 100$)

b — number of Sub-problem Selection Cards per subproblem ($1 \le b \le 2$)

K — number of Variable Format Cards ($1 \le K \le 10$)

Machine Requirements

STUDENTT requires 124K bytes during execution and uses a special OS Assembler Language program, BLNK, for missing data.

DESCRIPTIVE TITLE	Correlation and Multiple Linear Regression
CALLING NAME	UMST500
INSTALLATION NAME	University of Minnesota University Computer Center
AUTHOR(S) AND AFFILIATION(S)	Unknown
LANGUAGE	CDC Fortran IV
COMPUTER	CDC6600 (Scope 3.1.6)
PROGRAM AVAILABILITY	Deck and listing currently available
CONTACT	William Craig, EIN Tech. Rep., Center for Urban and Regional Affairs, Univ. of Minn., 311 Walter Library, Minneapolis, Minn. 55455 Tel.: (612) 373-7833

FUNCTIONAL ABSTRACT

This program performs a correlation analysis followed by a multiple linear regression. This program does *not* handle missing data. Any or all of the following may be calculated and given as output:

1. Raw cross-product matrix
2. Cross-product matrix about the means
3. Covariance matrix
4. Means and standard deviations
5. Correlation coefficient matrix
6. Multiple correlation coefficient
7. Ordinary regression coefficients and their standard errors
8. Normal regression coefficients and their standard errors
9. Total sum of squares of the dependent variable about its mean, separated into regression sum of squares and error sum of squares
10. Standard error of estimate
11. t statistics for testing significance of regression coefficients (regression coefficient divided by the standard error of that coefficient)
12. Partial correlation coefficients of the dependent variable with the independent variables

continued

13. Inverse of correlation matrix with dependent variable
14. Check on accuracy of inverse
15. Covariance matrix of normal regression coefficients
16. Inverse of cross-product matrix (about means) without the dependent variable
17. Back solution: observed dependent variable minus predicted dependent variable (residual) for all observations (The Durbin-Watson statistic is output when this option is chosen.)

Seven different transformations of all variables can be made. Twenty-one methods of transformation and/or generation of individual variables are possible. (See User Instructions: Problem Card and Transgeneration Cards.)

Restrictions on J, the number of variables:

$J \leq 216$ if only a correlation solution is called for; i.e., outputs 1–5 above.
$J \leq 145$ if a regression solution is called for, excluding outputs 14–16 above.
$J \leq 90$ if a regression solution is called for including all the above; i.e., outputs 1–17.

Restrictions on N, the number of observations (or individuals)

$N \leq 2^{16}-1 = 32{,}767$ if N is supplied to the program by an input card.
$N \leq 2^{47}-1 \doteq 10^{14}$ if the program is to count the number of observations as they are read in.

This program can compute any multiple regression equation which involves variables contained in the correlation matrix (output 5 above); 6–17 may be computed for each such equation. Any variable may be named as the dependent variable.

An optional feature of this program is the calculation of all possible regression equations involving *one dependent variable and up to 10 independent variables.* For example, if y is the dependent variable and x_1, x_2, x_3, are independent variables, then the following regressions would be obtained:

y and x_1
y and x_2
y and x_3
y and x_1, x_2
y and x_1, x_3
y and x_2, x_3
y and x_1, x_2, x_3

continued

An additional option is the use of a special calculation technique to compute the cross-product matrix (about means), without possible loss of significant digits, directly from the input data.

REFERENCES

Anderson, D. and Frisch, M., *Statistical Programs for Use on the CDC 6600 Computer* (Minneapolis, Minn.: Univ. of Minn., Univ. Comp. Ctr., 1969).

Welford, B.P., "Note on a Method for Calculating Corrected Sums of Squares and Products", Technometrics, *4* (1962), pp. 419-420.

Caffrey, J., "ALGORITHM 66", Communications of the Association for Computing Machinery, *4* (1961), p. 322.

Johnson, P. O., *Statistical Methods in Research* (New York: Prentice Hall, Inc., 1949), pp. 338-340.

A Multiple Regression and Correlation Program for the Univac Scientific (1103), RR–127 (New York: UNIVAC, 1956).

DESCRIPTIVE TITLE	Biomedical Computer Programs
CALLING NAME	BMD
INSTALLATION NAME	Washington University Computing Facilities
AUTHOR(S) AND AFFILIATION(S)	Health Sciences Computing Facility University of California, Los Angeles
LANGUAGE	FORTRAN IV
COMPUTER	IBM 360/50
PROGRAM AVAILABILITY	Proprietary; available for use at Washington University. Available for distribution from authors.
CONTACT	Dr. C. B. Drebes, Mgr., Scientific Data Processing, Computing Facilities, Box 1098, Washington University, St. Louis, Mo. 63130 Tel.: (314) 863-0100 ext. 3141

FUNCTIONAL ABSTRACT

The BMD package contains a variety of statistical programs in the following areas.

Description and Tabulation

Multivariate Analysis
a. Factor Analysis
b. Discriminant Analysis
c. Canonical Analysis

Regression Analysis
a. Linear
b. Polynomial
c. Asymptotic

Analysis of Variance and Covariance

Time Series Analysis

Special Programs
a. Life Table and Survival Rate
b. Contingency Table Analysis
c. Biological Assay
d. Guttman Scaling

continued

REFERENCES

Dixon, W.J., Ed., *BMD: Biomedical Computer Programs* (Berkeley: Univ. of Cal. Press, 1968).

Dixon, W.J., Ed., *BMD: Biomedical Computer Programs, X-series Supplement* (Berkeley: Univ. of Cal. Press, 1969).

DESCRIPTIVE TITLE	General Purpose Simulation System
CALLING NAME	GPSS
INSTALLATION NAME	Washington University Computing Facilities
AUTHOR(S) AND AFFILIATION(S)	IBM Application Program
LANGUAGE	360 Assembly Language
COMPUTER	IBM 360/50
PROGRAM AVAILABILITY	Proprietary; available for use at Washington University. Available for distribution from IBM.
CONTACT	Dr. C.B. Drebes, Mgr., Scientific Data Processing, Computing Facilities, Box 1098, Washington University, St. Louis, Mo. 63130 Tel.: (314) 863-0100 ext. 3141

FUNCTIONAL ABSTRACT

GPSS is a transaction-oriented language designed for conducting evaluations and experiments concerning the behavior of systems, methods and processes. It has a modular structure which permits "transactions" to flow through the system, where their interactions can be observed and modified. A "clock" is maintained by which events are either scheduled to occur or else determined by one of the eight random number generators provided. Information can be obtained regarding sequencing of operations, scheduling and allocation rules, inventories, queuing disciplines, machine failures, etc. In general, various trade-offs between cost and performance can be studied.

REFERENCES

General Purpose Simulation System/360, Introductory User's Manual (IBM Manual GH20-0304-0), (White Plains, N.Y.: IBM Corp. Tech. Pub. Depart., 1969).

General Purpose Simulation System/360, User's Manual (IBM Manual GH20-0326-3), (White Plains, N.Y.: IBM Corp. Tech. Pub. Depart., 1969).

General Purpose Simulation System/360, OS (360A-CS-17X) Operator's Manual (IBM Manual H20-0311), (White Plains, N.Y.: IBM Corp. Tech. Pub. Depart., 1969).

DESCRIPTIVE TITLE	Chi Square
CALLING NAME	UMST520
INSTALLATION NAME	University of Minnesota University Computer Center
AUTHOR(S) AND AFFILIATION(S)	Unknown
LANGUAGE	CDC Fortran IV
COMPUTER	CDC6600 (Scope 3.1.6.)
PROGRAM AVAILABILITY	Deck and listing presently available
CONTACT	William Craig, EIN Tech. Rep., Center for Urban and Regional Affairs, Univ. of Minn., 311 Walter Library, Minneapolis, Minn. 55455 Tel.: (612) 373-7833

FUNCTIONAL ABSTRACT

This program computes the chi-square criterion on two-way frequency tables. Input data consist of observations in a contingency table, commonly referred to as an r X c table, with the following restrictions:

$$r \geq 2, \quad c \geq 2$$

where r is the number of rows and c is the number of columns. Row and column sums and expected frequencies may be selected as an optional output. Yates' continuity correction is applied in the case of 2 X 2 tables.

Formulae

Let x_{ij} be the data element in the ith row and jth column.

Sums of columns: $n_{.j} = \sum_{i=1}^{r} x_{ij}$ where r is the number of rows.

Sums of rows: $n_{i.} = \sum_{j=1}^{c} x_{ij}$ where c is the number of columns.

continued

Total: $N = \sum_{j=1}^{c} n_{.j}$

Chi square:

$$\text{Chi square} = \sum_{i=1}^{r} \sum_{j=1}^{c} \frac{\left(x_{ij} - \frac{n_{.j}n_{i.}}{N}\right)^2}{\frac{n_{.j}n_{i.}}{N}}$$

$$= N \left[\sum_{i=1}^{r} \sum_{j=1}^{c} \left(\frac{x_{ij}^{2}}{n_{i.}\,n_{.j}}\right) - 1 \right]$$

The latter form is used in the program.

Degrees of freedom: d.f. = (r–1) (c–1)

Yates' continuity corrections:

Given the 2 X 2 table:

	a	b	a+b
	c	d	c+d
total	a+c	b+d	r

Use this table if $(a+c)(a+b)/r = a$

If $(a+c)(a+b)/r > a$, then use

$$\begin{array}{|cc|} \hline a + 1/2 & b - 1/2 \\ c - 1/2 & d + 1/2 \\ \hline \end{array} \quad \text{instead of} \quad \begin{array}{|cc|} \hline a & b \\ c & d \\ \hline \end{array}$$

If $(a+c)(a+b)/r < a$, then use

$$\begin{array}{|cc|} \hline a - 1/2 & b + 1/2 \\ c + 1/2 & d - 1/2 \\ \hline \end{array} \quad \text{instead of} \quad \begin{array}{|cc|} \hline a & b \\ c & d \\ \hline \end{array}$$

Expected frequencies: $f_{ij} = n_{.j}n_{i.}/N$

DESCRIPTIVE TITLE	Missing Data Correlation
CALLING NAME	UMST530
INSTALLATION NAME	University Computer Center University of Minnesota
AUTHOR(S) AND AFFILIATION(S)	Unknown
LANGUAGE	CDC COMPASS
COMPUTER	CDC 6600 (Scope 3.1.6)
PROGRAM AVAILABILITY	Deck and listing currently available
CONTACT	William Craig, EIN Technical Representative, Center for Urban and Regional Affairs, Univ. of Minn., 311 Walter Library, Minneapolis, Mn. 55455 Tel.: (612) 373-7833

FUNCTIONAL ABSTRACT

UMST530 computes Pearsonian product-moment correlations for all possible pairwise combinations of up to 130 variables. For this statistic and for the matrix of covariances, missing data for one or both variables of a pair exclude that observation from the analysis for that pair only; other pairs with non-missing data are retained. A count of non-missing observations for each set of pairs is kept and output. For each individual variable a count is also kept and output of the non-missing observations. For those observations a mean, variance and standard deviation are output.

Formulae

Let X_{ni} denote the nth observation of variable i where n = 1, 2, ..., NOBS and i = 1, 2, ..., NVAR, i.e. a total of NVAR variables with NOBS observations. Let $A_{nij} = 0$ if either X_{ni} or X_{nj} is missing and $A_{nij} = 1$ otherwise.

Count: $$N_{ij} = \sum_{n=1}^{NOBS} A_{nij}$$

continued

Mean: $\overline{X}_{ij} = \frac{1}{N_{ij}} \sum_{n=1}^{NOBS} A_{nij} X_{ni}$

Variance: $v_{ij} = \frac{1}{N_{ij}} \sum_{n=1}^{NOBS} A_{nij} (X_{ni} - \overline{X}_{ni})^2$

Covariance: $c_{ij} = \frac{1}{N_{ij}} \sum_{n=1}^{NOBS} A_{nij} (X_{ni} - \overline{X}_{ij})(X_{nj} - \overline{X}_{ji})$

Correlation: $r_{ij} = c_{ij} \ / \ (v_{ij} v_{ji})^{1/2}$

The user must note that $N_{ij} = 0$ is possible and that output results will be zero for that i and j, and mean nothing. Also, if $v_{ij} = 0$ (or $v_{ji} = 0$), the correlation will be zero and mean nothing.

DESCRIPTIVE TITLE	Nonparametric (Rank Order) Statistics
CALLING NAME	UMST540
INSTALLATION NAME	University of Minnesota University Computer Center
AUTHOR(S) AND AFFILIATION(S)	David Braddock Museum of Natural History Micheal Frisch University Computer Center
LANGUAGE	CDC Fortran IV
COMPUTER	CDC6600 (Scope 3.1.6)
PROGRAM AVAILABILITY	Deck and listing currently available
CONTACT	William Craig, EIN Technical Representative, Center for Urban and Regional Affairs, University of Minnesota, 311 Walter Library, Minneapolis, Minn. 55455 Tel.: (612) 373-7833

FUNCTIONAL ABSTRACT

This program is designed to rank input data, then calculate and print selected rank-order statistics. Any completely non-numeric (including blank) input field is considered to be missing data and calculations involving that field are not made. In making this check all non-numerics (except a leading minus) are stripped from the number, which will produce inaccurate results for the user who has non-aligned punched decimals, e.g. 0.8953 and 59.0 will be accepted as 8953 and 590, thus reversing the order of the numbers.

The problem size limitation is a function of the number of variables and the number of observations: 2 variables and 5000 observations or 100 variables and 100 observations both approach this limit. Statistics include:

1. Kruskal-Wallis One-Way Analysis of Variance both corrected for ties and uncorrected for *K* independent samples. Where appropriate, the chi-square probability is also printed. The ranking in this option is performed over all elements of data except for missing data.

continued

2. Spearman Rank Correlation between all pairs of variables over all observations where both variables represent non-missing data. The t statistic and the Student's t probability may also be output. R_s is fully corrected for ties.

3. Kendall Rank Correlation. All statements made in 2 above apply here except that the statistics printed are tau, Z, S and the normal distribution probability.

4. Kendall Coefficient of Concordance with its associated chi-square statistic and probability. As with 2 and 3 above, the ranking is performed within each separate variable, but here a single missing variable will delete an entire observation from the analysis.

Computation and Theory

The theory and computations of this program are to be found in Ref. 1.

Kruskal-Wallis One-Way Analysis of Variance

The statistic H is defined by the formula:

$$H(\text{uncorrected for ties}) = \frac{12}{N(N+1)} \sum_{j=1}^{k} \frac{R_j^2}{n_j} - 3(N+1)$$

where k = the number of variables

n_j = the number of observations for the jth variable (excluding missing observations)

$N = \sum_{j=1}^{k} n_j$, the total number of observations (excluding missing observations)

R_j = the sum of the ranks of observation for the jth variable

H corrected for ties is obtained by dividing the uncorrected H by

$$1 - \frac{\sum_{m=1}^{M} t_m^3 - t_m}{N^3 - N}$$

where t_m is the number of tied values in a tied group of observations and M is the number of tied groups.

Test of significance of H

Siegel[1] suggests that if not all $n_j \leq 5$, or $k \neq 3$, then a chi-

continued

square table may be used to test the significance of H(d.f. = k–1) (Table C in Siegel). In this case, the chi-square probability is printed by UMST540. Siegel's table 0 contains exact probabilities which can be used to test the significance of H when all $n_j \leq 5$ and $k = 3$.

Spearman Rank Correlation Coefficient for two variables, x and y

$$r_s = \frac{\Sigma x^2 + \Sigma y^2 - \Sigma d_i^2}{2\sqrt{\Sigma x^2 \ \Sigma y^2}}$$ this formula corrects for ties)

where

$$\Sigma x^2 = \frac{N^3 - N}{12} \quad T_x$$

where N is the number of observations, (missing data in one variable cause deletion of corresponding data in the other variable), and

$$T_x = \sum_{m=1}^{M_x} \frac{t_m^3 - t_m}{12}$$ for the x variable

t_m is the number of tied values in a tied group of observations and M_x is the number of tied groups in the x variable;

Σy^2 is the y-variable analog of x^2 ;

$\sum_{i=1}^{N} d_i^2$ is the numeric difference between rank on the x-variable and rank on the y-variable for observation i.

Test of significance of r_s

For large N (> 10, according to Siegel), the following statistic is distributed as Student's t with d.f. = N–2:

$$t = r_s \sqrt{\frac{N-2}{1-r_s^2}}$$

and a table of Student's t can be used to test the significance of this statistic. In this case, the Student's t probability

continued

is printed by UMST540. For smaller N, Siegel suggests that r_s be tested against an exact probability table. Table P in Siegel lists exact probabilities for $N \leq 30$.

Kendall Rank Correlation Coefficient (tau) for two variables, x and y

$$\text{Tau} = \frac{S}{\sqrt{1/2\ N(N-1) - T_x}\ \sqrt{1/2\ N(N-1) - T_y}}$$

where N is defined as for r_s above,

$$T_x = 1/2 \sum_{m=1}^{M_x} t_m\,(t_m - 1) \quad \text{for variable } x\ ,$$

T_y is the y-analog of T_x,

t_m, M_x and M_y are defined as for r_s above.

S is derived as follows: Let x be the vector of ranks of A, y be the vector of ranks of B. Place the elements of x in ascending order. Place the elements of y associated with those of x in the same order.

Then

$$\begin{aligned} S = {} & \text{SIGN}\,[y_2 - y_1] + \text{SIGN}\,[y_3 - y_1] + \dots + \text{SIGN}\,[y_n - y_1] \\ & + \text{SIGN}\,[y_3 - y_2] + \text{SIGN}\,[y_4 - y_2] + \dots + \text{SIGN}\,[y_n - y_2] \\ & + \dots + \dots + \text{SIGN}\,[y_n - y_{n-1}] \end{aligned}$$

where $\text{SIGN}\,[y_i - y_j] = 1$ if the expression in brackets is positive

$= -1$ if the expression in brackets is negative

$= 0$ if the expression in brackets is zero or if $x_i = x_j$

Test of Significance of Tau

For large N (Siegel suggests $N \geq 8$), Tau is distributed approximately as the normal distribution, and its significance may thus

continued

DESCRIPTIVE TITLE	Correlation and Orthogonal Factor Analysis
CALLING NAME	UMST550
INSTALLATION NAME	University of Minnesota University Computer Center
AUTHOR(S) AND AFFILIATION(S)	Lawrence Liddiard University Computer Center University of Minnesota
LANGUAGE	CDC Fortran IV
COMPUTER	CDC6600 (Scope 3.1.6)
PROGRAM AVAILABILITY	Deck and listing currently available
CONTACT	William Craig, EIN Tech. Rep., Center for Urban and Regional Affairs, Univ. of Minn., 316 Walter Library, Minneapolis, Minn. 55455 Tel.: (612) 376-7232

FUNCTIONAL ABSTRACT

This program performs correlation analysis followed by orthogonal factor analysis. The input is read into the computer using variable format; and the input data may have one of three forms: raw data by observation, lower triangle or a correlation (or pseudo-correlation) matrix, or orthogonal factor matrix. If the input is raw data, there are twenty-one possible methods of transformation and/or generation of individual raw variables. After the above methods, or separately, one of six complete transformations may be made on all of the variables (raw and generated).

The following is always given as printed output.

If raw data input:

> The standard error of a factor loading computed from the corrected formula given in Table B on page 441 of Ref. 1.
>
> A Name Card echo print if Name Cards are used.
>
> A Transgeneration Card echo print if transgeneration is used.
>
> Total number of observations.

continued

If raw data or correlation input:

The type of factor analysis component with the diagonal unchanged, common factor with squared multiple correlation on the diagonal, common factor with highest off diagonal absolute value on the diagonal, image covariance, image correlation or independent scale.

The average absolute value of the off-diagonal elements.

The trace (sum of elements on the diagonal) of the matrix to be factor analyzed.

The standard error of a factor loading times the square root of the number of observations.

The communality estimates (the diagonal of the matrix to be factor analyzed).

The eigenvalues for the principal-axis solution, the eigenvalue percent of the number of variables, accumulative percentage of the number of variables, the eigenvalue percent of the common variance (trace of the matrix), and the accumulative percentage of common variance for the computed number of factors.

For all input:

The Problem Card echo print
The Format Card(s) echo print
A message with respect to the total amount of computer core not used or needed above that alloted.

Any or all of the following may be calculated and given as output:

1. Means, standard deviations, variances, covariances, correlations and raw cross products (not usually chosen as output because of massiveness).
2. Means, standard deviations and the elements of the lower triangle of the correlation matrix.
3. An echo of the input correlation or pseudo-correlation matrix.
4. The means and standard deviations on a separate output file for factor scoring or punch output.
5. The elements of the lower triangle of the image analysis matrix (correlation, covariance or independent scale).
6. The data-to-data transformation matrix of image analysis on a separate output file for factor scoring or punch output.
7. The principal-axis factor-loading matrix for the computed number of factors; bordered on the left by the names

continued

(if supplied), indices and communalities (sum of squared row elements) of the variables; bordered on the bottom by the variances (sum of squared column elements) and percentages of trace (variance*100/trace) of the factors.

8. The sorted principal-axis factor-loading matrix for the computed number of factors; bordered on the bottom by the variances and percentages of trace of the factors. Each column of the factor loading matrix is sorted in descending order with the corresponding variable indices or names given in a column to the left of that factor column.
9. The principal-axis factor-scoring matrix for the computed number of factors; bordered on the left by the names or indices and a line count within a variable set. This same information is also put on a separate output file for factor scoring or punch output.
10. The quartimax factor-loading matrices for a stepped range of factors and bordered as in 7.
11. The sorted quartimax factor-loading matrices for a stepped range of factors and bordered as in 8.
12. The quartimax factor-scoring matrices for a stepped range of factors and bordered as in 9. This same information is also put on a separate output file for factor scoring or punch output.
13. The varimax factor-loading matrices for a stepped range of factors and bordered as in 7.
14. The sorted varimax factor-loading matrices for a stepped range of factors and bordered as in 8.
15. The varimax factor-scoring matrices for a stepped range of factors and bordered as in 9. This same information is placed on a separate output file for factor scoring or punch output.

While UMST550 does not compute factor scores itself, auxillary programs are available which can compute and print those scores for each observation when raw data are input. A user wishing to compute factor scores must store his raw data and relevant intermediate results on scratch tape (TAPE99). When filling out the Problem Card, the user must select those options which put that data onto TAPE99. If a user desires factor score output, a note explaining the methodology a user desires to follow in computing those factor scores must accompany the user's EIN submission.

Limits inherent to UMST550

Total number of variables, $2 \leq NV \leq 250$ (approximately) for a CDC 6600 with a 65K-word core.
Total number of observations if given on Problem Card $2 \leq NOBS \leq 99999$

continued

Number of Variable Format Cards, $1 \le NFC \le 9$
Number of Transgeneration Cards $0 \le NTGC \le 999$
Number of factors $1 \le NF \le NV/3$ unless a specific stepped range is chosen or forced in Cols. 59 and 60 of the Problem Card.

Limits which may be supplied by the user

Lower limit for factor eigenvalue (the principal factor variance)— a value of 1.0 may be supplied by the program
Orthogonal rotation cycle limit—a maximum of 40 or NV/2 cycle limit may be supplied by the program
Orthogonal rotation radian angle limit—.005 radians may be supplied by the program

REFERENCES

1. Harman, H.H., *Modern Factor Analysis* (Chicago: Univ. of Chicago Press, 1960).

2. Anderson, D. and Frisch, M., *Statistical Programs for Use on the CDC 6600 Computer* (Minneapolis, Minn.: Univ. of Minn., Univ. Comp. Ctr., 1969).

3. Caffrey, J., "Invrs" Algorithm 66, *Comm. ACM* (July 1961), p. 322.

4. Horst, P., *Factor Analysis of Data Matrices* (New York: Holt, Rinehart and Winston, Inc., 1965).

5. Hotchkiss, R., "SORT2", a standard sorting routine at the Univ. Comp. Ctr. at the Univ. of Minn. based on T.H. Hibbard's "An Empirical Study of Minimal Storage Sorting," *Comm. ACM, 6,* (1968) pp. 206-212.

6. Ortega, J., "The Givens Householder Method for Symmetric Matrices," in A. Ralston and H. Wilf, *Mathematical Methods for Digital Computers, II,* (New York: John Wiley & Sons, Inc., 1967), pp. 94–115.

7. Wilkinson, J.H., "Householder's Method for Symmetric Matrices," Numerische Mathematik, *4* (1962), pp. 354–361.

8. Wilkinson, J.H., "Calculation of the Eigenvalues of a Symmetric Matrix by the Method of Bisection," Numerische Mathematik, *4* (1962), pp. 362–367.

9. Wilkinson, J.H., "Calculation of the Eigenvectors of a Symmetric Matrix by Inverse Iteration," Numerische Mathematik, *4* (1962), pp. 368–376.

DESCRIPTIVE TITLE	Least Square Curve Fitting Using Orthogonal Polynomials
CALLING NAME	UMST560
INSTALLATION NAME	University of Minnesota University Computer Center
AUTHOR(S) AND AFFILIATION(S)	Unknown
LANGUAGE	CDC Fortran IV
COMPUTER	CDC6600 (Scope 3.1.6)
PROGRAM AVAILABILITY	Deck and listing currently available
CONTACT	William Craig, EIN Tech. Rep., Center for Urban and Regional Affairs, Univ. of Minn., 311 Walter Library, Minneapolis, Minn. 55455 Tel.: (612) 373-7833

FUNCTIONAL ABSTRACT

Using a modified method of least squares, this program fits polynomials to pairs of input data vectors. It first calculates a polynomial of degree 1, then of degree 2, and so on up to the degree of the polynomial specified.

Output and Options

Different weights for each data element may be specified.
After every iteration the error sum of squares and standard error of the coefficients are printed.
After every iteration observed values for both X and Y, predicted Y, observed minus predicted Y, and weights assigned to each point may be requested. In this notation X is the independent variable and Y the dependent variable. This option will be called the Full Output. It is useful information when the degree of the polynomial desired is known.
An automatic termination with the polynomial of best fit may be selected.
The coefficients of the orthogonal polynomials used in the calculation of the least squares polynomials may be printed.
The input data may be transformed and vectors added according to specifications supplied to the program. Pairs of X–Y vectors are selected for curve fitting if there are more than two variables. By use of transformation 04,

continued

$x' = e^x$, an exponential fit of the data may be obtained of the form $y = a_0 + a_1e^x + a_2e^{2x} + \ldots$

A graph of the polynomial together with observed values may be printed.

Method

Let X_i denote the independent variable (abscissa), Y_i the dependent variable (ordinate), W_i the corresponding weight ($W_i = 1$ if weights are not supplied), and $P_N(X) = \sum_{i=0}^{N} B_iX^i$ the least-square polynomial of degree N. If $P_J(X)$ denotes any Jth-degree polynomial, then $P_N(X)$ is that polynomial which minimizes $\sum_{i=1}^{M} W_i[Y_i - P_J(X_i)]^2$.

Because of the much greater accuracy attainable, orthogonal polynomials, rather than the standard matrix inversion technique, are used to determine $P_N(X)$. There is no restriction on the spacing of the X_i. Calculation and printout begin with degree 1 and proceed to the highest degree selected.

If the automatic termination option is selected, the program terminates when the "error sum of squares", a standard criterion of computational accuracy, is not reduced for two consecutive higher degrees. The full output and differences are then printed out for that polynomial which resulted in the smallest "error sum of squares". This process constructively defines the polynomial of "best fit" since the probability is small that polynomials of higher degree will yield a smaller "error sum of squares."

Equations

$$M_W = \sum_{i=1}^{M} W_i \text{ ; } M = \text{number of data points}$$

$$M_N = \text{maximum of } (1, M-N-1) \text{ ; } N = \text{degree of polynomial}$$

$$P_N(X) = \sum_{i=1}^{M} B_iX^i \text{ ; least squares polynomial}$$

continued

Means:

a. $\overline{X} = \sum_{i=1}^{M} W_i X_i / M_W$

b. $\overline{Y} = \sum_{i=1}^{M} W_i Y_i / M_W$

Total sum of squares:

$$Y_s = \sum_{i=1}^{M} W_i Y_i{}^2 - \overline{Y}^2 \cdot M_W$$

Standard deviations:

a. $S_X = [\sum_{i=1}^{M} W_i X_i{}^2 - \overline{X}^2 \cdot M_W]/(M_W - 1)$

b. $S_Y = Y_S/(M_W - 1)$

Error sum of squares:

$$E = \sum_{i=1}^{M} W_i [P_N(X_i) - Y_i]^2$$

Standard Error of Y Estimate:

$$S_E = \sqrt{E/M_N}$$

Note: S is undefined when $N = M-1$. In this case $M_N = 1$ is used in the calculation.

Standard Error of the Coefficients:

$SB_i = \sqrt{(\text{Variance } B_i)/M_N}$ See Ref. 1 for computation of these. See also previous note on M_N.

Back solution:

$P_N(X_i)$ = Predicted Y ; Observed Y − Predicted Y = $Y_i - P_N(X_i)$

continued

Time Estimate

The maximum time a program will run is a function of the number of observations, the number of variables, the number of Select Cards, and the number of graphs printed. The equation given below assumes a graph size of 50 by 100 lines and a Full Output option.

$$\text{TIME} = [.15\ \text{NVAR} + \text{NSELE} * \text{NOBS}^{.38} + \text{NOBS}\ (.00036) + 4.62 * \text{NGRAF}]\ \text{Seconds}$$

Where: NSELE = the number of Select Cards
NVAR = the number of variables
NOBS = the number of observations
NGRAF = the number of 50 X 100 graphs printed out.

REFERENCES

1. Todd, J., *A Survey of Numerical Analysis* (New York: McGraw-Hill Book Co., Inc., 1962), pp. 347–362.

2. Hildebrand, F.B., *Introduction to Numerical Analysis* (New York: McGraw-Hill Book Co., Inc., 1956) pp. 258–287.

3. Jordan, T.L., and Vogel, R.E., "Least Squares Polynomial Fitting," *Systems Programs for CO-OP*, E2 UCSD LSQPOL, (Palo Alto, Calif: Control Data Corp.).

4. Hamming, R.W., *Numerical Methods for Scientists and Engineers* (New York: McGraw-Hill Book Co., Inc. 1962).

DESCRIPTIVE TITLE	Multivariate Analysis of Variance
CALLING NAME	UMST570
INSTALLATION NAME	University of Minnesota University Computer Center
AUTHOR(S) AND AFFILIATION(S)	William Walster, David Doren University Computer Center University of Minnesota
LANGUAGE	CDC Fortran IV
COMPUTER	CDC6600 (Scope 3.1.6)
PROGRAM AVAILABILITY	Deck and listing currently available
CONTACT	William Craig, EIN Tech. Rep., Center for Urban and Regional Affairs, Univ. of Minn., 311 Walter Library, Minneapolis, Minn. 55455 Tel.: (612) 373-7833

FUNCTIONAL ABSTRACT

This program performs multivariate analysis of variance and/or covariance on a vector of dependent variables.

Input

The input data can be from one to many dependent variates and/or covariates. Each observation *must* have a complete set of scores on all the variates and covariates. It may have unequal observations in the cells of the selected design.

Options

Linear and/or eleven algebraic transformations can be made when data are read in.

The analyses of all anova designs are based on contrasts (comparisons). The program generates several main effect options for between analyses or allows freedom for the programmer to supply his own design matrix (set of contrasts). The program forces orthogonality among contrasts. All the above contrasts are for "between-cells analyses".

continued

Sources of Variation

a. The program automatically groups the appropriate sets of contrast estimates into main effects and appropriate interaction effects. Options are available to regroup the contrast estimates. This allows tests on specific contrasts or contrast sets other than the normal main effects and interaction effects.

b. *The order of the contrasts in the design matrix are of a specific form.* It is advised that the user acquaint himself with the section on Regrouping Cards to fully understand the "conventional" order utilized by the program. This "conventional" order may be changed by the user through the "reorder" option (see section on Special Orderings). The order of the tests of the sets of contrasts takes on special meaning since each test assumes the previous test was null.

A test of the equality of the variance-covariance matrices between cells can be performed. When there is only one dependent variate, this test is the same as the test of equality of variances.

Either weighted or unweighted means analysis may be chosen independently for each factor in the design.

A second linear transformation of variates can be performed to allow orthogonal linear combinations of original variates within each observation. The resulting analysis, utilizing proper recombinations, can give a multivariate analog to univariate repeated-measures analyses.

The recombination option allows the design analysis to be performed on selected subsets of variates (and covariates).

DESCRIPTIVE TITLE	Stepwise Regression
CALLING NAME	UMST580
INSTALLATION NAME	University of Minnesota University Computer Center
AUTHOR(S) AND AFFILIATION(S)	Mr. M. Dale Fimple Sandia Corporation
LANGUAGE	CDC Fortran IV Modified for use at Univ. of Minnesota
COMPUTER	CDC6600 (Scope 3.1.6)
PROGRAM AVAILABILITY	Deck and listing currently available
CONTACT	William Craig, EIN Tech. Rep., Center for Urban and Regional Affairs, Univ. of Minn., 311 Walter Library, Minneapolis, Minn. 55455 Tel.: (612) 373-7833

FUNCTIONAL ABSTRACT

This program performs stepwise linear regression handling as many as 80 independent and 25 dependent variables. There is an option for using variable weights on the observations. Thirty-four methods of transformation and/or generation of individual variables as possible. This program selects a subset of independent variables by fixed F or fixed probability, deleting least significant variables one at a time.

Output includes regression and correlation parameters at each step. An optional feature of this program can compute and give as output residual information. Since the data analyzed may differ from the input data, all or part of the data analyzed may also be included as output.

Any or all of the following may be calculated and given as output:

1. Sums, sums of squares, and sums of cross products
2. Means, sums of squares, and sums of cross products
3. Correlation coefficients
4. The inverse of the correlation matrix and the inverse of the X'X matrix.

continued

Method

The method used by UMST580 to pick the most significant subset of independent variables for each dependent variable is as follows:

1. The regression analysis is performed on the entire set of P independent variables.
2. For each variable in the analysis, the statistic, b_i^2/C_{ii}, is computed where b_i is the regression coefficient giving the relationship between the ith independent variable and the dependent variable. C_{ii} is the ith diagonal element of $(X'X)^{-1}$. This statistic gives the reduction in the regression sum of squares when X_i is deleted from the analysis.
3. The minimum of b_i^2/C_{ii} is obtained.
4. This value is divided by the error mean square at that point, and this ratio has an F distribution.
5. This value of F is tested against either a value of F or a probability level included in the control card for the problem.
6. If it is determined that this minimum sum of squares is significant, the deletion process is discontinued. Otherwise, X_i is deleted, the inverse matrix is adjusted for the deletion, and a branch is taken back to step 2.

Since the method given above for selecting the optimum subset of variables is equivalent to partially reinverting the X'X matrix, it is essential to include a check on the accuracy of the inverse. It should be pointed out that all operations are carried out on the correlation matrix which has all elements between $\pm$ 1. The basic inversion routine used throws out any variables which would cause singularity. After the correlation matrix is inverted, the norm of $(I - RB_0)$ is computed. I is the $P \times P$ identity matrix, R is the correlation matrix, and B_0 is the computed estimate of R^{-1}. The norm used is defined by

$$N(A) = \sqrt{\underset{ij}{\Sigma\Sigma} a^2_{ij}}$$

If this norm does not meet the requirements specified on the control card, the option is available to use Hotelling's method[1] to obtain a better estimate of the inverse. This is an iterative technique, and the ith estimate of the inverse is given by

$$B_i = B_{i-1}(2I - RB_{i-1}).$$

continued

If the norm mentioned above is less than 1.0, convergence to the true inverse is assured (in theory). Actually, the degree of convergence is restricted by the use of floating point arithmetic.

The following table[2] gives probability values computed by UMST580 for selected F ratios with one degree of freedom in the numerator and the number indicated in the denominator. The true probability values are given across the top.

Degress of Freedom	True Probability 0.70	0.90	0.95	0.99
1	0.6927	0.8687	0.9073	0.9386
2	0.6988	0.8950	0.9422	0.9806
3	0.6991	0.8992	0.9477	0.9869
4	0.7001	0.9005	0.9497	0.9887
5	0.7000	0.9014	0.9505	0.9895
6	0.7000	0.9020	0.9510	0.9897
7	0.6989	0 9022	0.9512	0.9899
8	0.6997	0.9025	0.9515	0.9902
9	0.6994	0.9026	0.9517	0.9903
10	0.7003	0.9025	0.9517	0.9901
12	0.6986	0.9031	0.9520	0.9903
14	0.6996	0.9030	0.9520	0.9903
16	0.6998	0.9033	0.9520	0.9903
18	0.6996	0.9035	0.9521	0.9903
20	0.6989	0.9032	0.9522	0.9903
40	0.6965	0.9041	0.9524	0.9903
100	0.6981	0.9043	0.9528	0.9903
500	0.7002	0.9045	0.9527	0.9903

REFERENCES

1. Hotelling, H., "Some New Methods in Matrix Calculation," Annals of Math. Stat., *14* (1943), pp. 1—34.

2. Kendall, M.G., *The Advanced Theory of Statistics* (2nd ed.; London: Griffin Pub. Co., 1967), *II*, Equation 21.62, p. 118.

000 0145

DESCRIPTIVE TITLE	Cross Classification
CALLING NAME	UMST590
INSTALLATION NAME	University of Minnesota University Computer Center
AUTHOR(S) AND AFFILIATION(S)	Robert Ellis, Department of Journalism University of South Dakota
LANGUAGE	CDC Fortran IV
COMPUTER	CDC6600 (Scope 3.1.6)
PROGRAM AVAILABILITY	Deck and listing currently available
CONTACT	William Craig, EIN Tech. Rep., Center for Urban and Regional Affairs, Univ. of Minn., 311 Walter Library, Minneapolis, Minn. 55455 Tel.: (612) 373-7833

FUNCTIONAL ABSTRACT

This program generates a cross-tabulation, i.e. a bivariate frequency distribution, for pairs of variables. As an option, a distribution may be formed for each value of a third variable or for each pair of values of a third and a fourth variable, i.e., using one or two control variables. Missing data options may interpret blanks and non-numerics as missing data, or read missing data tags, or convert plus and minus punches to numeric values. The data may be used as given, or they may be grouped into intervals before a plot is formed. The data may be real or integer numbers. A transgeneration subprogram is included.

The following may be selected as output.

1. Chi square and contingency coefficient C.
2. Column and row percentages and percentages to the overall total of the table.
3. Nonparametric correlation coefficients: Kendall tau with associated z; Goodman–Kruskal gamma, and Somers dyx and dxy. (Note: for a 2X2 table, gamma = Yule's Q.)

Limitations

1. Number of variables $2 \leq NVAR \leq 80$
   Number of observations $NOBS \leq 40,000$

continued

2. Total data input (NOBS X NV) $\leq$ 40,000 where NV is the largest integer contained in $\frac{\overline{\text{N}}\text{VAR} + 9}{10}$;
   e.g., 10,000 observations on 40 variables, or
   5,000 observations on 80 variables.
3. The number of levels on any variable must be $\leq$ 20, not including missing data. Values taken on may be any real number, but only 20 distinct values may occur. In order to meet this restriction, Maximum and Minimum Cards, transgeneration or one of the recoding options (Interval or PICK) may be used.
4. Maximum table size is 20 X 20.

Formulae

Let x_{ij} be the data element in the ith row and jth column.

Sums of columns:

$$n_{.j} = \sum_{i=1}^{r} X_{ij} \text{ where r is the number of rows.}$$

Sums of rows:

$$n_{i.} = \sum_{j=1}^{c} X_{ij} \text{ where c is the number of columns.}$$

Total:

$$N = \sum_{j=1}^{c} n_{.j}$$

Chi square:

$$\text{Chi square} = \sum_{i=1}^{r} \sum_{j=1}^{c} \left[\frac{[X_{ij} - \frac{n_{.j}n_{i.}}{N}]^2}{\frac{n_{.j}n_{i.}}{N}} \right] = N\left[\left\langle \sum_{i=1}^{r} \sum_{j=1}^{c} \frac{X_{ij}{}^2}{n_{i.}n_{.j}} \right\rangle - 1\right]$$

The latter form is used in the program.

Degrees of freedom: d.f. = (r–1)(c–1)
Where missing data or subtables result in zero rows or columns, the message 'ADJUSTED TO ______ FOR THIS TABLE' will appear.

Yates' continuity correction, applied to 2 X 2 tables:

If $X_{ij} < f_{ij}$ then use $X_{ij} + 1/2$

continued

If $X_{ij} > f_{ij}$ then use $X_{ij} - 1/2$

If $X_{ij} = f_{ij}$ then use X_{ij}

Expected frequencies:

$$f_{ij} = n_{i.} \text{ X } n_{.j}/N$$

Contingency coefficient:

$$C = \sqrt{\text{Chi square} / (N + \text{Chi square})}$$

where Chi square has not been corrected for continuity in the case of a 2 X 2 table.

$$Q = \sum_{i=1}^{r-1} \sum_{j=2}^{c} (X_{ij}X_{kl}) \quad \text{where } k = (i+1), \ldots, r \text{ and } l = 1, \ldots, (j-1)$$

$$P = \sum_{i=1}^{r-1} \sum_{j=1}^{c-1} (X_{ij}X_{kl}) \quad \text{where } k = (i+1), \ldots, r \text{ and } l = (j+1), \ldots, c$$

$$S = P - Q$$

Kendall Tau:

1. For square table (i.e., r = c):

$$\text{Tau B} = \frac{S}{\sqrt{1/2N(N-1)-T} \quad \sqrt{1/2N(N-1)-U}}$$

$$\text{where } T = \frac{1}{2} \sum_{i=1}^{r} n_{i.}(n_{i.} - 1)$$

$$\text{and } U = \frac{1}{2} \sum_{j=1}^{c} n_{.j}(n_{.j} - 1)$$

2. For non-square table (i.e., r ≠ c):

$$\text{Tau C} = \frac{2S}{N^2(\frac{m-1}{m})} \quad \text{where m is smaller of r, c}$$

continued

Sampling variance of S:

$$\text{Var} = \frac{1}{18}\left[N(N-1)(2N+S) - \sum_{i=1}^{r} n_{i.}(n_{i.}-1)(2n_{i.}+5) - \sum_{j=1}^{c} n_{.j}(n_{.j}-1)(2n_{.j}+5)\right]$$

$$+ \frac{1}{9N(N-1)(N-2)}\left[\sum_{i=1}^{r} n_{i.}(n_{i.}-1)(n_{i.}-2)\right]$$

$$\left[\sum_{j=1}^{c} n_{.j}(n_{.j}-1)(n_{.j}-2)\right] + \frac{1}{2N(N-1)}\left[\sum_{i=1}^{r} n_{i.}(n_{i.}-1)\right]$$

$$\left[\sum_{j=1}^{c} n_{.j}(n_{.j}-1)\right]$$

Normal deviate:

$$Z(s) = \frac{S}{\sqrt{\text{Var}}}$$

(Continuity correction): Given a 2 X 2 table the absolute value of S is reduced by 1/2 N with the restriction that if $|S| \leq 1/2\ N$, $S = 0$. For all other tables, $|S|$ is reduced by 1.

Goodman-Kruskal Gamma:

$$\text{Gamma} = S/(P+Q)$$

Somers d:

$$d_{xy} = \frac{S}{\frac{1}{2}\left[N^2 - (\sum_{i=1}^{r} n^2_{i.})\right]} \qquad d_{xy} = \frac{S}{\frac{1}{2}\left[N^2 - (\sum_{j=1}^{c} n^2_{.j})\right]}$$

continued

REFERENCES

Hays, W.L., *Statistics for Psychologists* (New York: Holt, Rinehart & Winston Inc., 1963), pp. 652-656.

Kendall, M.G., *Rank Correlation Methods* (3rd ed.; New York: Hafner Pub. Co., 1962).

McNemar, Q., *Psychological Statistics* (3rd ed.; New York: John Wiley & Sons, Inc., 1962).

Somers, R.H., "A New Asymmetric Measure of Association for Ordinal Variables," Am. Soc. Rev., *27* (1962), pp. 799–816.

DESCRIPTIVE TITLE	Descriptive Statistics
CALLING NAME	UMST600
INSTALLATION NAME	University of Minnesota University Computer Center
AUTHOR(S) AND AFFILIATION(S)	Unknown
LANGUAGE	CDC Fortran IV
COMPUTER	CDC6600 (Scope 3.1.6)
PROGRAM AVAILABILITY	Decks and listings currently available
CONTACT	William Craig, EIN Tech. Rep. Center for Urban and Regional Affairs, Univ. of Minn., 311 Walter Library, Minneapolis, Minn. 55455 Tel.: (612) 373-7833

FUNCTIONAL ABSTRACT

UMST600 provides a nearly complete description of the distribution of each of up to 999 variables. For each variable (X) the following measures are calculated: $\sum X$, $\sum X^2$, $\sum(X-\bar{X})^2$, mean, variance, and standard deviation. Each measure is produced after reading in all observations, but options exist to identify missing data so that the measures may pertain to a different subset of the total number of observations for each variable. The count of non-missing observations is output for each variable. There is no limit on the number of observations to be processed.

Several options are available which make the program more useful. The user may opt to produce $\sum X^3$, $\sum X^4$, $\sum(X-\bar{X})^3$ and $\sum(X-\bar{X})^4$. He may also opt to print the maximum and minimum value for each variable. X value (from largest to smallest), frequency, cumulative frequency, cumulative proportion and absolute proportion for each variable may be output. Choosing this option limits the size of the job that can be processed in one use of the program to 400 variables and 100 observations or 10 variables and 4000 observations. Variables may also be named, attaching an eight-character label to each variable's output.

continued

Notation and Computations

N = the number of observations

Variance formula 1

$$s^2 = \frac{N\,\Sigma\, X^2 - (\Sigma\, X)^2}{N(N-1)} \quad \text{(unbiased)}$$

Variance formula 2

$$s^2 = \frac{N\,\Sigma\, X^2 - (\Sigma\, X)^2}{N^2} \quad \text{(biased)}$$

$$\Sigma(X - \overline{X})^2 = \Sigma X^2 - N\overline{X}^2$$

$$\Sigma(X - \overline{X})^3 = \Sigma X^3 - 3\overline{X}\Sigma X^2 + 2N\overline{X}^3$$

$$\Sigma(X - \overline{X})^4 = \Sigma X^4 - 4\overline{X}\Sigma X^3 + 6\overline{X}^2\ \Sigma X^2 - 3N\overline{X}^4$$

DESCRIPTIVE TITLE	Chi Square from Raw Data
CALLING NAME	UMST620
INSTALLATION NAME	University of Minnesota University Computer Center
AUTHOR(S) AND AFFILIATION(S)	Unknown
LANGUAGE	CDC Fortran IV
COMPUTER	CDC6600 (Scope 3.1.6)
PROGRAM AVAILABILITY	Decks and listing currently available
CONTACT	William Craig, EIN. Tech. Rep., Center for Urban and Regional Affairs, Univ. of Minn., 311 Walter Library, Minneapolis, Minn. 55455 Tel.: (612) 373-7833

FUNCTIONAL ABSTRACT

This program computes and prints a frequency distribution of two variables based on intervals supplied to the program.

Optional outputs include:

1. A check of all the frequencies as calculated above to determine if all are equal to or above the given frequency supplied to the program. If cell frequencies are below this value, they are printed out and the problem terminated.
2. The chi-square statistic with its associated degrees of freedom.
3. Marginal sums of the calculated frequencies.
4. Expected frequencies based on the marginal sums and the total frequency.

Input data consist of observations to be grouped in a contingency table, commonly referred to as an r X c table with $r \geq 2$ and $c \geq 2$. Yates' continuity correction is applied in the case of 2 X 2 tables.

continued

Formulae

Let x_{ij} be the data element in the ith row and jth column.

Sums of columns: $n_{.j} = \sum_{i=1}^{r} x_{ij}$ where r is the number of rows.

Sums of rows: $n_{i.} = \sum_{j=1}^{c} x_{ij}$ where c is the number of columns.

Total: $N = \sum_{j=1}^{c} n_{.j}$

Chi square:

$$\text{Chi square} = \sum_{i=1}^{r} \sum_{j=1}^{c} \frac{\left(x_{ij} - \frac{n_{.j}n_{i.}}{N}\right)^2}{\frac{n_{.j}n_{i.}}{N}}$$

$$= N\left[\left(\sum_{i=1}^{r} \sum_{j=1}^{c} \frac{x_{ij}^{2}}{n_{i.}n_{.j}}\right) - 1\right]$$

The latter form is used in the program.

Degrees of freedom: d.f. = (r–1) (c–1)

Yates' continuity corrections:

Given the 2 X 2 table:

	a	b	a+b
	c	d	c+d
total	a+c	b+d	r

Use this table if $(a+c)(a+b)/r = a$

If $(a+c)(a+b)/r > a$, then use

a + 1/2	b – 1/2
c – 1/2	d + 1/2

instead of

a	b
c	d

continued

If $(a+c)(a+b)/r < a$, then use

$$\boxed{\begin{matrix} a - 1/2 & b + 1/2 \\ c + 1/2 & d - 1/2 \end{matrix}} \quad \text{instead of} \quad \boxed{\begin{matrix} a & b \\ c & d \end{matrix}}$$

Expected frequencies: $f_{ij} = n_{.j}n_{i.}/N$

DESCRIPTIVE TITLE	General Linear Hypothesis for Anova, Unequal Frequencies
CALLING NAME	UMST610
INSTALLATION NAME	University of Minnesota University Computer Center
AUTHOR(S) AND AFFILIATION(S)	Modified version of BMD05V from Health Sciences Comp. Fac., UCLA
LANGUAGE	CDC Fortran IV
COMPUTER	CDC6600 (Scope 3.1.6)
PROGRAM AVAILABILITY	Deck and listing currently available
CONTACT	William Craig, EIN Tech. Rep., Center for Urban and Regional Affairs, Univ. of Minn., 311 Walter Lib., Minneapolis, Minn. 55455 Tel.: (612) 376-7232

FUNCTIONAL ABSTRACT

This program (a modification of UCLA's program BMD05V) analyzes the statistical significance of dependent variables for those experimental designs that can be formulated in terms of the General Linear Hypothesis model. It allows processing of the same design for as many dependent variables as field length will allow, where all variables are read in at once and the calculations are performed on the variables one at a time.

This program does *not* allow analysis of covariance. It does *not* allow transformation or transgeneration. The program can analyze unbalanced analysis of variance designs according to stated hypotheses.

Description of Model

Suppose we have n observations (values assumed by n random variables) $y_1, y_2, \dots, y_n$, which have the linear structure

$$y_\alpha = x_{1\alpha}\beta_1 + x_{2\alpha}\beta_2 + \dots + x_{\ell\alpha}\beta_\ell + e_\alpha,$$

$$\alpha = 1, 2, \dots, n$$

continued

where $\{x_{i\alpha}\}$ are known constants, $\{\beta_i\}$ are unknown variables, and $\{e_\alpha\}$ are uncorrelated random variables, with means 0 and variances σ^2.

The program is designed so that

$$x_{ij} = x_{i2}, x_{i2}, \ldots, x_{in}$$

is the design or analysis of variance type. The analysis of variance parameters are to be represented by the numbers 0, 1, −1 (or actually −9 to +9).

Using this program it is possible to process a great variety of analysis of variance designs; for example, full factorials, partial factorials, unbalanced designs, and experimental designs with missing values.

For a given set $y_\alpha, x_{1\alpha}, \ldots, x_{\ell\alpha}$ the program computes the least squares estimates,

$$\hat{\beta}_i, \quad i = 1, 2, \ldots, \ell,$$

the residual sum of squares

$$\sum_{\alpha=1}^{n} (y_\alpha - \sum_{i=1}^{\ell} x_{i\alpha}\hat{\beta}_i)^2,$$

and the degrees of freedom for this sum of squares. Hypotheses of the form,

$$H: \beta_{i_1} = \beta_{i_2} = \ldots = \beta_{i_k} = 0,$$

where $(i_1, i_2, \ldots, i_k)$ is any subset of $(1, 2, \ldots, \ell)$ may be stated. Least-squares estimates of the β's under these hypotheses are computed along with the residual sums of squares, and the degrees of freedom for these sums.

If we now make the assumption that the e_α are normally distributed, then the above information is sufficient to compute the usual F tests for the hypotheses stated, with the proper degrees of freedom.

DESCRIPTIVE TITLE	Single and Simultaneous Equation (TSLS, LISE) Regression Package
CALLING NAME	UMST630
INSTALLATION NAME	University of Minnesota University Computer Center
AUTHOR(S) AND AFFILIATION(S)	Morris Norman University of California, Santa Barbara James Weatherby University of Texas
LANGUAGE	CDC Fortran IV
COMPUTER	CDC6600 (Scope 3.1.6)
PROGRAM AVAILABILITY	Deck and listing currently available
CONTACT	William Craig, EIN Tech. Rep., Center for Urban and Regional Affairs, Univ. of Minn., 311 Walter Library, Minneapolis, Minn. 55455 Tel.: (612) 373-7833

FUNCTIONAL ABSTRACT

This is a multipurpose program designed to estimate the coefficients of a multiple regression model or a simultaneous equation model. Three estimation techniques are available to the user: ordinary least squares (OLS)[1,2]; two stage least squares (TSLS)[3,4]; and limited information maximum likelihood (LISE)[5,6]. The program has variable input format and a transformation routine with nine options including a provision to lag variables one period. For the computational methods used see Ref. 7.

The standard printout for each equation estimated is as follows:

a. Augmented moment matrix (raw cross-product matrix)
b. Variance-covariance matrix of the coefficients
c. Regression coefficients and their asymptotic standard errors
d. Variance and standard deviation of the equation
e. R^2 adjusted for degrees of freedom
f. R^2 not adjusted for degrees of freedom
g. t test for each regression coefficient; H_0: $B = 0$, H_1: $B \neq 0$
h. F test; H_0: $(B_1,\ldots,B_n) = 0$ H_1: $(B_1,\ldots,B_n) \neq 0$.

continued

The optional printout for each equation estimated is as follows:

a. Covariance and correlation matrix for all of the variables
b. Graph of actual and predicted value of the dependent variable versus the corresponding observation
c. Durbin-Watson statistic
d. Scatter diagrams of values of the dependent variable plotted against the values of selected independent variable(s)
e. Elasticities of each independent variable with respect to the dependent variable (taken about the means)
f. Standard partial regression coefficient for each independent variable
g. If the transformation routine is used, a table describing the indicated transformation is printed. The number of transformations is restricted to forty.

Formulae of Selected Statistics

Moment Matrix of all the variables (optional)

$$= \frac{\sum_{k=1}^{n} Z_{ik}Z_{jk} - \frac{n\,\overline{Z}_i\,\overline{Z}_j}{(n-1)}}{n}$$

Correlation Matrix of all the variables (optional)

$$= \frac{\text{Moment Matrix}}{(\sigma_i \, X \, \sigma_j)}$$

Variance—Covariance Matrix of Coefficient[8]

$$S_{cb,cb} = S^2 \frac{Y_1' \, X(X'X)^{-1}X'Y_1 \quad Y_1'X_1}{X_1'Y_1 \qquad\qquad X_1'X_1}$$

This is the analogue of E(β − Eβ)(β − Eβ). The standard errors given are the square roots of the diagonal elements of the matrix (these are given in parentheses).

Unadjusted R^2

$$R^2 = 1 - \frac{\sum_{t=1}^{n} (y_t - \hat{y}_t)^2}{\sum_{t=1}^{n} (y_t - \overline{y})^2}$$

continued

Adjusted R^2

$$\tilde{R}^2 = 1 - \frac{\dfrac{\sum_{t=1}^{n} (y_t - \hat{y}_t)^2}{n - m - 1}}{\dfrac{\sum_{t=1}^{n} (y_t - \bar{y})^2}{n - 1}}$$

$$\beta_j = \frac{\hat{\beta}_j \ SE_j}{\sqrt{\dfrac{\sum_{t=1}^{n} (y_t - \hat{y}_t)^2}{n - m - 1}}}$$

Elasticity of y_t with respect to X_j

$$X_j = \frac{\bar{X}_j}{\bar{y}} \hat{\beta}_j$$

Symbols:

Z_j = the dependent or independent variable

y_t = dependent variable

$\hat{y}_t$ = estimated value of y_t

$\bar{y}$ = mean of y_t

n = number of observations

m = number of independent variables

$\hat{\beta}_j$ = estimate of β_j, the regression coefficient of the jth independent variable

SE_j = the standard error of $\hat{\beta}_j$

$\bar{X}_j$ = the mean of the jth independent variable

continued

REFERENCES

1. Goldberger, A.S., *Econometric Theory* (New York: John Wiley & Sons, Inc., 1964), Ch. 4.

2. Johnston, J., *Econometric Methods* (New York: McGraw Hill, Inc., 1963), Ch. 4.

3. Goldberger, pp. 329–336.

4. Johnston, pp. 258–260.

5. Goldberger, pp. 338–345.

6. Johnston, pp. 254–258.

7. Christ, C.F., *Econometric Models and Methods* (New York: John Wiley & Sons, Inc., 1966), Ch. 9, Sec. 9–10.

8. Goldberger, p. 333.

DESCRIPTIVE TITLE	University Computing Facility
INSTALLATION NAME	Marquette University Computing Center
LANGUAGES	FORTRAN III, FORTRAN IV, COBOL, MAP
COMPUTER	IBM 7040 with 16K-word core storage
AVAILABILITY	Open shop
CONTACT	Mr. Robert Miller, Acting Director, Computing Center, Marquette University, 1515 W. Wisconsin Ave., Milwaukee, Wisc. 53233 Tel.: (414) 224-7700

FUNCTIONAL ABSTRACT

The IBM 7040 at Marquette University is offered to EIN users with Computer Center personnel available for consultation in utilizing the Program Library and in correcting user-written programs.

Equipment at the Marquette University Computing Center includes the following.

Controllers

IBM 1414 Model I	Synchronizer for tapes
IBM 1414 Model IV	Synchronizer for cards, printer

Storage

IBM 729	Tape drives (2)
IBM 7330	Tape drives (4)

I/O

IBM 1402	Card reader-punch
IBM 1403	Line printer

Peripheral equipment (not connected to the 7040)

IBM 407	Tabulating machine
IBM 514	Reproducing punch
IBM 083	Sorter
IBM 870	Graph plotter
IBM 056	Verifier
IBM 1232	Optical reader

continued

Peripheral equipment (continued)

IBM 534	Card punch
IBM 026	Printing card punches(8)
GDDR 3B	Gerber data reader
Microdyne	Analog-to-digital converter

Magnetic tape is available on a loan basis in 200-, 1200-, and 2400-foot sizes. Programs available at Marquette include those developed at the Computing Center and programs ordered through SHARE, with an ordering time of two to three weeks for complete documentation on the latter.

DESCRIPTIVE TITLE	Kiewit Computation Facility
INSTALLATION NAME	Kiewit Computation Center Dartmouth College
LANGUAGES	Dartmouth BASIC Dartmouth FORTRAN Dartmouth ALGOL Dartmouth LISP LAFFF (Language to Aid Financial Fact Finders) CRIII (Computer Research Involving Investment Information) MIX GEFORT (GE FORTRAN) GMAP (GE Macro Assembly Program) TRAC (Text Reckoning and Compiling Program)
COMPUTER	GE-635
AVAILABILITY	Available to specified categories of remote terminals
CONTACT	For Account Initiation, Period of Usage, and Policy: A. Kent Morton, EIN Technical Representative, Kiewit Computation Center, Dartmouth College, Hanover, N.H. 03755 Tel.: (603) 646-2864

For specific problems in the following areas (all extensions preceded by Area Code 603, exchange 646-):

Area	*Name*	*Extension*
Validating user numbers and passwords; granting special permissions	Nancy Broadhead Stephen V.F. Waite	2643 2643
Programming assistance	Diane Mather Stephen V.F. Waite	3283 2643
Use of Plotters	Arthur Luehrmann	2976,2864
Supplying Manuals	Jann Dalton	2147

continued

FUNCTIONAL ABSTRACT

The Dartmouth Time Sharing System is a remote access facility supporting such low-speed terminals as Teletype models 33, 35, and 37, Friden 7100 and IBM 2741. Detailed terminal, coupler, and bit rate specifications are contained in references 1 and 2. On-site peripherals are not available to remote users, since the Center does not handle input or output beyond the operations counter.

Hardware available to remote users includes the dual processor GE 635 with 160K x 36-bit words; two GE Datanet 30's with a total of 155 bit buffer units; one GE magnetic drum with controllers; two IBM 2314 disc units with a Datametrics interface to a GE IOC controller; and six tape drives (7 track) operated by two tape controllers.

The executive system was written and is maintained and modified by Dartmouth undergraduates under the supervision of the Software Development Director at Kiewit.

Remote faculty users will be alloted 16K of core memory, 64 seconds of run-time, and, in most cases, 6K of disc storage.

Special software available to all users includes the LAFFF and the CRIII systems developed by the Amos Tuck School of Business to aid in finding financial facts (see Ref. 3); a simulator of the MIX system designed by Donald Knuth; RUNOFF (for text processing); Dartmouth EDIT, TEXTEDIT, and STRING EDITOR, all for file manipulation, whether the file be a program, text, or alphanumeric data; TEACH, which provides the possibility for writing a program to automatically test programs written by students for a given problem or course; and IMPRESS, which consists of numerous sociological data files and a core of programs for manipulating them.

REFERENCES

1. Hargraves, R., and Mather, D., *User's Guide to the Dartmouth Time-Sharing System*, Dartmouth College TM 022 (Hanover, N.H.: Oct. 1970).

2. Morton, A.K., *Hardware and Access Supplement for Remote Users* (Dart. Col., Jan. 1971).

3. *CRIII* (Tuck Sch. of Bus. Admin., Dart. Col., 1970). CRIII is a system for financial fact finding written in BASIC. It is now being used almost exclusively with and by Tuck School students, and is obtainable from the Tuck School.

DESCRIPTIVE TITLE	Remote Job Entry System (RJE)
INSTALLATION NAME	The Pennsylvania State University Computation Center
LANGUAGE	Full facilities of the IBM Operating System for the 360.
COMPUTER	IBM 360/67
AVAILABILITY	Remote access
CONTACT	Dr. Daniel Bernitt, The Pennsylvania State University Computation Center, 105 Computer Building, The Pennsylvania State University, University Park, Pa. 16802 Tel.: (814) 865-9527

FUNCTIONAL ABSTRACT

The Remote Job Entry (RJE) system permits research users to access the System 360/67 computer from a remote keyboard terminal in a time-sharing mode. The terminals supported by the system are IBM 2741's with BCD or "SELECTRIC" character set, DATEL or other terminals that simulate the 2741 and hence have the same characteristics, IBM 1050 terminals with the BCD character set, and Teletypes.

Principal Equipment

Main Campus

Number	*Type*	*Use*
1	IBM 360/67	Main system
5	IBM 360/20	High speed entry
15	IBM 1050	Low speed entry
18	IBM 2741	
20	DATEL Thirty-21	
1	IBM 2250	Interactive Graphics
1	CalComp Plotter (30 inch bed)	Hard Copy Graphics

Branch Campuses

Number	*Type*	*Use*
10	IBM 2780	High speed entry
1	IBM 360/20	High speed entry
3	IBM 1050	Low speed entry

continued

Storage and I/O Devices

The 360/70 central system consists of: 1 million bytes main memory; 2 million bytes LCS; six IBM 2400 tape drives (9-track); two IBM 2400-1 tape drives (7-track); two IBM 2301 Drums; three IBM 2314 Disk Drives: one IBM 2703 Communications Adapter; two IBM 2540 Card reader/punches; three IBM 2403 Printers; one IBM 2671 Paper Tape reader; five IBM 2260 CRT.

Software

1. Complete OS/360 (currently Release 18 + HASP).
2. Locally developed remote job entry, supporting low speed terminals of the type

   IBM 2741 (EBCDIC or SELECTRIC character set) or
   DATEL THIRTY-21 (portable, similar to 2741's) or
   IBM 1050 (EBCDIC character set).
3. Various language processors and application programs. Includes systems developed locally and at other independent computing facilities along with all IBM type I and II systems.

REFERENCES

The Pennsylvania State University Computation Center, *Remote Job Entry System: General Information and Summary Description* (University Park, Pa.: The Penn. State Univ., Nov. 1970). Available from the EIN Office at the cost of reproduction and mailing.

DESCRIPTIVE TITLE	Vogelback Computing Facility
INSTALLATION NAME	Vogelback Computing Center Northwestern University
LANGUAGES	FTN (CDC FORTRAN Extended) RUN (CDC FORTRAN) COBOL, SNOBOL, LISP, ALGOL, MIMIC, SLIP, SIMSCRIPT, COMPASS
COMPUTER	CDC 6400, 65K
AVAILABILITY	Batch processing, remote batch processing, remote on-line processing
CONTACT	Lorraine Borman, EIN Technical Representative, Vogelback Computing Center, Northwestern University, 2129 Sheridan Road, Evanston, Ill. 60201 Tel.: (312) 492-3682

FUNCTIONAL ABSTRACT

The 6400 is primarily a batch processing machine but can handle on-line access through Teletype models 33 or 35, or a CDC 200 User Terminal.

Equipment

1	6416 Central computer (65K memory)
1	6603 disk
2	854 disk pack drives
2	501 printers (upper case only)
1	415 card punch
2	405 card readers
1	565 CalComp plotter
4	607 magnetic tape transports (7 track)
1	3691 paper tape reader/punch
1	6671 communications multiplexor

The Center is currently running under the SCOPE 3.2 operating system. Data tapes must be 7 track, BCD; only 026 character set is acceptable on either punched cards or tapes.

Use of the Center is restricted to educational and non-profit institutions.

DESCRIPTIVE TITLE	UCSB On-Line System
CALLING NAME	OLS
INSTALLATION NAME	University of California Santa Barbara UCSB Computer Center
AUTHOR(S) AND AFFILIATION(S)	Dr. David Harris, Principal Investigator Ron Stoughton, Principal Programmer UCSB Computer Research Lab.
LANGUAGE	Basic Assembler Language
COMPUTER	IBM System 360/75
PROGRAM AVAILABILITY	Listing of programs may be obtained by special arrangement with the Computer Research Lab.
CONTACT	Glenn Davis, OLS Manager, Computer Center, Univ. of Cal., Santa Barbara, Cal. 93106 Tel.: (805) 961-2462

FUNCTIONAL ABSTRACT

The UCSB On-Line System (OLS) provides the capability for sophisticated mathematical analysis for use in solving problems where human interaction is either necessary or desired.

OLS accepts both real and complex numbers (scalars) as operands as well as lists of such numbers (vectors). Operations performed on scalars produce scalar results, which can be numerically displayed; operations on vectors produce vector results (the specified operation being performed on each component), and results of computation can be displayed either numerically or graphically. Operands can be stored and used as required. Operators include sine, cosine, logarithm, and exponentiation; and each is executed with a single button push. Facility is provided for interaction between operands of different types (e.g. vectors and scalars). In addition, a limited set of operations manipulate integers used in subscripting.

Additional features are provided to support OLS's basic mathematical capability. Although OLS normally executes each

continued

button as it is pushed, a button sequence may be defined, named, and saved for later execution. Convenient means are provided for editing such sequences. Lists of buttons to be executed can include programmed pauses, allowing manual and programmed activity to be interfaced; as well as branching based upon results of computation. Messages can be composed of alphameric, Greek, and special characters, and displayed. Those characters not specifically provided by OLS may be designed by the user and stored, and then are available for use. A collection of button lists and user-created characters is referred to as a "system". Systems are named and can be permanently stored and later retrieved. Portions of systems may be transferred between systems, and systems may be transferred between users. Sets of scalars and vectors may also be named, permanently stored, and later retrieved.

Apart from OLS's mathematical capability, a recent development provides the ability to create and edit a "deck" of cards and submit it for execution in an OS partition. Operations on string, record and file levels are provided. Data-sets residing on any disk pack within the installation may be fetched, examined, modified, and submitted for execution. Work continues in the general field and further developments are expected.

DESCRIPTIVE TITLE	MERC Time-Sharing System
INSTALLATION NAME	Middle-Atlantic Educational and Research Center
LANGUAGES	FORTRAN IV, COBOL, ASSEMBLY, BASIC, and various special applications programs.
COMPUTER	RCA Spectra 70/46G with 256 kilobytes of core
AVAILABILITY	Remote and mail access
CONTACT	Dr. Richard S. Lehman, Director MERC, P. O. Box 1372, Lancaster, Pa. 17604 Tel.: (717) 393-0132 or 393-5021 Dr. Paul W. Ross, Manager of Services, MERC, P. O. Box 1372, Lancaster, Pa. 17604 Tel.: (717) 393-0132 or 393-5021

FUNCTIONAL ABSTRACT

MERC is making available its RCA Spectra 70/46G with 2-megabyte drum, four disc storage units and four tape drives (9-channel, 800 bpi). This system is a highly interactive time-sharing and simultaneous batch computer operating system. The terminals supported are Model 33 and 35 teletypewriters, via 1/2 state WATS lines in the 215 and 717 dialing areas. Local dial-up in the Lancaster area and other selected locations in eastern Pennsylvania and southern New Jersey is available.

The system supports programs in BASIC, FORTRAN IV, COBOL, and Assembly. The FORTRAN system has both background compilation and execution, fast batch and interactive FORTRAN. COBOL is supported for conventional background batch-mode operation or in a mode in which input may be taken from a terminal. In addition, a syntax checking system is available.

Entry of data is via paper tape reader at the user's site, or by input of card decks at the computer center in Lancaster, for building large data base files. In addition, the system provides a large data base management system known as QWIK-TREIVE for special on-line data management applications. Various other special application packages are available.

continued

REFERENCES

Coia, E.J., *Interactive BASIC System Reference Manual* (Lancaster, Pa.: MERC, July 1970).

MERC Program Library Manual (Lancaster, Pa.: MERC, July 1970).

Gallagher, J., *MERC Calculator Manual* (Lancaster, Pa.: MERC, July 1970).

Copies of these manuals are available from MERC or through EIN at the cost of reproduction and mailing.

DESCRIPTIVE TITLE	Catholic University Computer Center
INSTALLATION NAME	The Catholic University of America Computer Center
LANGUAGES	FORTRAN IV BASIC COBOL ASSEMBLER (MACRO) and various special languages and utilities
COMPUTER	PDP-10 with 64K words of core
AVAILABILITY	Remote access teletype terminals and Batch processing
CONTACT	Dr. Andrew G. Favret, Dir., Computer Center, The Catholic Univ. of America, Washington, D.C. 20017 Tel.: (202) 529-6000 ext. 661

FUNCTIONAL ABSTRACT

The PDP-10 is a timesharing system with a rather flexible Monitor. The equipment includes: two 7-track tape drives, one RPO2 disk pack drive, eight DECtapes, and one CalComp on-line graphic plotter, in addition to card reader, line printer and paper-tape reader/punch. All of the peripherals are available directly or indirectly to remote users. Remote access low-speed terminals such as Teletype models 33 and 35 should be used (ASCII Code). The dial-up number is (202) 526-3300. It is necessary to obtain project/programmer numbers and password before using the system.

I
N
D
E
X

DESCRIPTIVE-TITLE INDEX

EIN NO.

A

000 0077	Algorithm for the Optimization of a Quadratic Form Subject to Linear Restraints, An (ZORILLA)
000 0076	Algorithms for Analysis of Variance and Covariance of Incomplete Block and Lattice Designs (GAVIAL)
000 0062	Analysis of Change-over Experiments (ZFE-03; ZFE-04)
000 0089	—— of Contingency Tables (ACT version 1.00)
000 0051(p)	——-of-Variance Method of Unweighted Means (ANOVUM)
000 0108	—— of Variance Using Between-Subject Designs (SABCA)
000 0051(r)	—— of Variance with Repeated Measures (Proportionality Assumed) (AOVRM)
000 0051(q)	—— of Variance with Unequal Subclass Numbers (Method of Expected Subclass Numbers) (ANOVES)
000 0016	Automatic Interaction Detector (AID)

B

000 0051(s)	Bartlett's Test for Homogeneity of Variance (BARTL)
000 0059	Basic Information Retrieval System (BIRS)
000 0053	BEEF Data Processing Subroutines (BEEFDP)
000 0054	—— Mathametical Subroutines (BEEFM)
000 0114	Best Least Squares Polynomial Approximation Subroutine (LSPOL)
000 0132	Bio-Medical Computer Programs (BMD)
000 0055	—— Multivariate Statistical Programs (BMD)
000 0018	Biserial—Point-Biserial Correlation Program (BISR)
000 0006	"Book-Type" Indexing Program (INDEXER)
000 0084	Bucharest Sort a List into Ascending Order (ABSRT)
000 0085	—— Sort a List into Descending Order (DBSRT)

C

000 0052	CalComp Plotter Subroutines
000 0019	Canonical Correlation (CANON)
000 0051(g)	—— Correlation (CANON)
000 0136	Catholic University Computer Center
000 0137	Chi Square (UMST520)
000 0020	—— and C Statistics (CHISQR)

continued

EIN NO.	
000 0147	Chi Square from Raw Data (UMST620)
000 0051	Collection of Statistical Programs (STPAC)
000 0123	Comparison of Characteristics from "Ethnographic Atlas" (ETHATLAS)
000 0044	Computer Means, Median, Frequencies (PROFILE)
000 0134	Continuous System Modeling Program (CSMP)
000 0130	Correlation and Multiple Linear Regression (UMST500)
000 0141	—— and Orthogonal Factor Analysis (UMST550)
000 0029	Critical-Path Summary—Probability Statistic Based on Normal Curve (PERTC)
000 0145	Cross Classification (UMST590)
000 0118	Cross-Cultural Comparison (POLYCOMP)
000 0124	Cultural Comparison (CULTCOMP)
000 0125	Culture Identification (CULTPIK)
	D
000 0117	Data Survey and Normality Test (NORMSURV)
000 0122	Demographic Information from Murdock's *Ethnographic Atlas* (ETH—INFO)
000 0146	Descriptive Statistics (UMST600)
000 0021	Discriminant Analysis for Two Groups (DISCRIM)
	E
000 0022	Eigensystems of Nonsymmetric, Real Square Matrices (EIGSYS)
000 0073	Equipercentile Equating Program (SCORMACH)
	F
000 0002	Factor Analysis (FACTOR)
000 0065	—— Analysis by Direct "Oblimin" Method (OBLIMIN)
000 0051(i)	—— Analysis (or Principal-Components Analysis) (FANAL)
000 0127	Finished Characteristics from Textor's *A Cross-Cultural Summary* (TEX—CODE)
000 0126	—— Characteristics of 400 Cultures (TEXTOR)
000 0128	Fitting Nonlinear Regression Functions (TARSIER)
000 0050(a)	FORMAC Utility Program (FMACUT)

continued

EIN NO.	
000 0140	FORTRAN Computer Based Serials Holding Management (UPDATE/HOLD/PUBLSH/LANSUB)
000 0042	—— Program for Computing a Full Set of Canonical Correlations Relating Two Sets of Measurements (CANON)
000 0007	—— Program for Evaluating Internal Consistency and Single-Factoredness in Sets of Multilevel Attitude Items (ATTANAL)
000 0009	—— Program for Guttman and Other Scalogram Analyses (GUTTSCL)
000 0072	—— Program to Assist in the Process of Political Reapportionment (BELOW)
000 0082	—— Subroutine Package to Solve Ordinary Differential Equations (NODE)
000 0083	—— Subroutine Package to Solve Ordinary Differential Equations (DNODE)
000 0086	—— Subroutine to Solve Simultaneous Linear Equations with Complex Coefficients (CGELG)
000 0051(a)	Frequency Analysis with Chi Square (FAWCS)
000 0081	FURNIVAL Regression Screen (FURNIVAL)
	G
000 0148	General Linear Hypothesis for Anova, Unequal Frequencies (UMST610)
000 0107	—— Multiple Regression Analysis (REGAN1)
000 0116	—— Program for Multivariate Cross-Classification (NUCROS)
000 0133	—— Purpose Simulation System (GPSS)
000 0008	Generalized Analysis of Transaction Flows (INDIFF)
000 0017	—— Stepwise Regression (B34T)
000 0098	—— t Test (UOM 5)
	H
	I
000 0106	Individual Case Statistics Program (SUMSCRDS)
000 0024	Information-Oriented Language—A Generalized Information and Retrieval System (INFOL)
000 0099	—— Retrieval Services
000 0034	—— Retrieval System for Creating, Maintaining, Indexing, and Retrieving from Files of Textual Information (TRIAL)

continued

EIN NO.	
000 0058	Inquirer II System for Content Analysis (I/II)
000 0033	Interrupt Time Series—Three Tests of Significance (TIMEX)
000 0131	Invert Ill-behaved Matrices Using Hotelling and Bodewig's Iterative Technique (HRDMIN)

J

K

000 0051(1)	Kendall Rank Correlation Coefficients (tau) (KETAU)
000 0049	Key Grapheme in Context (KGIC)
000 0101	Kiewit Computing Facility (Dartmouth)
000 0051(o)	Kruskal—Wallis One-Way Analysis of Variance (KRWAL)

L

000 0142	Least Square Curve Fitting Using Orthogonal Polynomials (UMST560)
000 0025	Linear Programming (LINPROG)

M

000 0119	Major Characteristics from Murdock's *Ethnographic Atlas* (ETH—CODE)
000 0001	Mann—Whitney U Test (MANNWH)
000 0051(n)	———Whitney U Test (MNWHT)
000 0060	Matching Factor Solutions (MATCHFS)
000 0093	Mathematical Programming System (MPS)
000 0061	Matrix Decomposition for Points of View Analysis (MATDEC)
000 0091	—— Operations (AES 106)
000 0104	MERC Time-Sharing System
000 0092	Mineral Identification (MINERAL)
000 0087	Minres Method of Factor Analysis (MINRES)
000 0138	Missing Data Correlation (UMST530)
000 0070	Multidimensional Scalogram Analysis (MSA-I)
000 0012	Multiple Regression and Correlation Analysis (BMD29)
000 0011	—— Regression and Correlation Analysis, Modified with Plots (PLOTYY)
000 0074	—— Scalogram Analysis (MSA)
000 0026	Multivariate Analysis of Variance (MANOVA)

continued

EIN NO.	
000 0063	Multivariate Analysis of Variance (MANOVA)
000 0143	—— Analysis of Variance (UMST570)
000 0005	—— Analysis of Variance-Covariance, Hotelling's T (MULCVR)
	N
000 0027	95X95 Factor Analysis with Varimax Rotation (MESA1)
000 0068	Nonmetric Multidimensional Scaling (KRUSKAL)
000 0075	—— Multidimensional Scaling (KRUSCAL)
000 0139	Nonparametric (Rank Order) Statistics (UMST540)
000 0014	Northwestern University Correlation Analysis (NUCORR)
000 0028	—— University Cross Classification and Tabulation (NUCROS)
000 0109	Numerical Frequency Analysis (NUMFREQ)
	O
000 0038	O'Neill Concordance Package (CONTEXT, CONCORD)
	P
000 0041	Package of Three FORTRAN Programs for Computerized Assistance in the Instruction of Beginning and Remedial Reading and the Evaluation of Such Instruction, A (DOVACK)
000 0112	Page Plotter Using the Line Printer (LPLOT)
000 0051(f)	Partial Correlation (PARCOR)
000 0051(d)	Pearson Product—Moment Correlation Coefficient (PPMCR)
000 0051(k)	Phi Coefficient (PHICO)
000 0050	PL/I—FORMAC Interpreter (FORMAC)
000 0069	Plot (PLOT)
000 0135	Primal—Dual Transportation Algorithm, A (TRANSPRT)
000 0023	Program Deck Identification Field Sequencing Program (IDSEQN)
000 0043	Program for Analysis of Linear Systems (PALS)
000 0030	Punch and Print Rankings (PPRANK)

continued

EIN NO.	
	Q
000 0048	Quick-Draw Graphics System (QDGS)
	R
000 0121	Random Sample of 50 Cultures (ETH—RAND)
000 0120	—— Samples of Cultures (ETH—DGRE)
000 0031	Rank-Order Statistics (RKSTAT)
000 0067	Rao Constellation and Distance Analysis (DISCRIM2)
000 0088	Reformat into Either BCD or EBCDIC (EDP C 005 E; EDP C 005 B)
000 0078	Regression Model Building System (MOUFLON)
000 0102	Remote Job Entry System (RJE)
	S
000 0032	Sequence Checking (SEQCHK)
000 0051(e)	Significance of Pearson Product—Moment Correlation Coefficient (SIGPP)
000 0110	Simple Multiple Linear Regression (MISREGN)
000 0071	Simulation Package for University Research and Training (SPURT)
000 0037	Simulator of SAMOS (A Simple Imaginary Machine Language for Instruction) (SAMOS)
000 0149	Single and Simultaneous Equation (TSLS, LISE) Regression Package (UMST630)
000 0113	Smallest Space Analysis (SSA-1)
000 0051(m)	Spearman Rank Correlation Coefficients (rho) (SPRHO)
000 0096	—— Rank-Order Correlation (UOM 4)
000 0046	Static Leontief Input—Output Analysis (INOUT)
000 0045	Statistical Analysis of Single Equation Models (QSASE)
000 0051(b)	—— Summary (STSUM)
000 0051(h)	Step-Up and Step-Down Multiple Linear Regression (UPREG/DNREG)
000 0010	Stepwise Multiple-Discriminant Analysis (EIDISC)
000 0144	—— Regression (UMST580)
000 0115	—— Regression Analysis (STEPREGN)
000 0064	Structure-Factor Least-Squares Refinement Program for IBM 7090 (ORFLS-PX)
000 0090	Student Scheduling System (SCHEDULE)
000 0003	Student's T and F Ratio (STDNF2)

continued

EIN NO.

000 0013 Student's T Test (STDNF1)
000 0129 —— t Test (STUDENTT)
000 0004 Subtle, Unbiased, Zealous Yatagen of Questionnaires (SUZYQ)
000 0039 Synagraphic Computer-Mapping Program (SYMAP)
000 0047 —— Mapping (SYMAP)
000 0080 —— Mapping Program (SYMAP)

T

000 0035 Test Scorer and Statistical Analysis (TSSA)
000 0111 ——-Scoring and Item-Analysis Package (QUICKSCORE/ITEMSTEP/ITEMRS)
000 0079 TEXT360: A System for Producing Manuals (TEXT360)
000 0040 Transportation Planning Package (TRAN/PLAN)
000 0051(c) T Test of Difference between Means (TTEST)

U

000 0105 UCSB On-Line System (OLS)
000 0056 UNIVAC 1107 Linear Programming Package (LP1107)
000 0057 —— 1107 PERT/COST, PERT/TIME System (PERT)
000 0094 Univariate and Multivariate Analysis of Variance, Covariance and Regression (NYBMUL)
000 0100 University Computing Facility (Marquette)
000 0095 —— of Maryland Test Scoring Program (Version 4) (UOM 32)
000 0066 Unrestricted Maximum Likelihood Factor Analysis (UMLFA)
000 0015 Upper- and Lower-Case Greek and Roman Alphabetic Plotting (SCRIPT)

V

000 0051(j) Varimax Rotation (VARMX)
000 0103 Vogelback Computing Facility

W

000 0097 Weighted Summing Test Scoring Program (Recoding) (UOM 87)

continued

EIN NO.	
	X, Y
	Z
000 0036	Zellner's Three-Stage Least-Squares Program (Z3SLS)

EIN NUMBER INDEX

000 0001	Mann—Whitney U Test (MANNWH)
000 0002	Factor Analysis (FACTOR)
000 0003	Student's T and F Ratio (STDNTF2)
000 0004	Subtle, Unbiased, Zealous Yatagen of Questionnaires (SUZYQ)
000 0005	Multivariate Analysis of Variance-Covariates, Hotelling's T (MULCVR)
000 0006	"Book-Type" Indexing Program (INDEXER)
000 0007	FORTRAN Program for Evaluating Internal Consistency and Single-Factoredness in Sets of Multilevel Attitude Items (ATTANAL)
000 0008	Generalized Analysis of Transaction Flows (INDIFF)
000 0009	FORTRAN Program for Guttman and Other Scalogram Analyses (GUTTSCL)
000 0010	Stepwise Multiple-Discriminant Analysis (EIDISC)
000 0011	Multiple Regression and Correlation Analysis, Modified with Plots (PLOTYY)
000 0012	Multiple Regression and Correlation Analysis (BMD29)
000 0013	Student's T Test (STDNTF1)
000 0014	Northwestern University Correlation Analysis (NUCORR)
000 0015	Upper- and Lower-Case Greek and Roman Alphabetic Plotting (SCRIPT)
000 0016	Automatic Interaction Detector (AID)
000 0017	Generalized Stepwise Regression (B34T)
000 0018	Biserial—Point-Biserial Correlation Program (BISR)
000 0019	Canonical Correlation (CANON)
000 0020	X^2 and C Statistics (CHISQR)
000 0021	Discriminant Analysis for Two Groups (DISCRIM)
000 0022	Eigensystems of Nonsymmetric, Real Square Matrices (EIGSYS)
000 0023	Program Deck Identification Field Sequencing Program (IDSEQN)
000 0024	Information-Oriented Language—A Generalized Information and Retrieval System (INFOL)
000 0025	Linear Programming (LINPROG)
000 0026	Multivariate Analysis of Variance (MANOVA)
000 0027	95X95 Factor Analysis with Varimax Rotation (MESA1)
000 0028	Northwestern University Cross Classification and Tabulation (NUCROS)
000 0029	Critical-Path Summary—Probability Statistic Based on Normal Curve (PERTC)
000 0030	Punch and Print Rankings (PPRANK)
000 0031	Rank-Order Statistics (RKSTAT)
000 0032	Sequence Checking (SEQCHK)

continued

000 0033	Interrupt Time Series—Three Tests of Significance (TIMEX)
000 0034	Information Retrieval System for Creating, Maintaining, Indexing, and Retrieving from Files of Textual Information (TRIAL)
000 0035	Test Scorer and Statistical Analysis (TSSA)
000 0036	Zellner's Three- Stage Least-Squares Program (Z3SLS)
000 0037	Simulator of SAMOS (A Simple Imaginary Machine Language for Instruction) (SAMOS)
000 0038	O'Neill Concordance Package (CONTEXT, CONCORD)
000 0039	Synagraphic Computer-Mapping Program (SYMAP)
000 0040	Transportation Planning Package (TRAN/PLAN)
000 0041	A Package of Three FORTRAN Programs for Computerized Assistance in the Instruction of Beginning and Remedial Reading and the Evaluation of Such Instruction (DOVACK)
000 0042	FORTRAN Program for Computing a Full Set of Canonical Correlations Relating Two Sets of Measurements (CANON)
000 0043	Program for Analysis of Linear Systems (PALS)
000 0044	Computer Means, Median, Frequencies (PROFILE)
000 0045	Statistical Analysis of Single Equation Models (QSASE)
000 0046	Static Leontief Input—Output Analysis (INOUT)
000 0047	Synagraphic Mapping (SYMAP)
000 0048	Quick-Draw Graphics System (QDGS)
000 0049	Key Grapheme in Context (KGIC)
000 0050	PL/I—FORMAC Interpreter (FORMAC)
000 0050(a)	FORMAC Utility Program (FMACUT)
000 0051	Collection of Statistical Routines (STPAC)
000 0051(a)	Frequency Analysis with Chi Square (FAWCS)
000 0051(b)	Statistical Summary (STSUM)
000 0051(c)	T Test on Difference between Means (TTEST)
000 0051(d)	Pearson Product—Moment Correlation Coefficient (PPMCR)
000 0051(e)	Significance of Pearson Product—Moment Correlation Coefficient (SIGPP)
000 0051(f)	Partial Correlation (PARCOR)
000 0051(g)	Canonical Correlation (CANON)
000 0051(h)	Step-Up and Step-Down Multiple Linear Regression (UPREG/DNREG)
000 0051(i)	Factor Analysis (or Principal-Components Analysis) (FANAL)
000 0051(j)	Varimax Rotation (VARMX)
000 0051(k)	Phi Coefficient (PHICO)
000 0051(l)	Kendall Rank Correlation Coefficients (tau) (KETAU)
000 0051(m)	Spearman Rank Correlation Coefficients (rho) (SPRHO)
000 0051(n)	Mann—Whitney U Test (MNWHT)

continued

000 0051(o)	Kruskal—Wallis One-Way Analysis of Variance (KRWAL)
000 0051(p)	Analysis-of-Variance Method of Unweighted Means (ANOVUM)
000 0051(q)	Analysis of Variance with Unequal Subclass Numbers (Method of Expected Subclass Numbers) (ANOVES)
000 0051(r)	Analysis of Variance with Repeated Measures (Proportionality Assumed) (AOVRM)
000 0051(s)	Bartlett's Test for Homogeneity of Variance (BARTL)
000 0052	CalComp Plotter Subroutines
000 0053	BEEF Data Processing Subroutines (BEEFDP)
000 0054	BEEF Mathematical Subroutines (BEEFM)
000 0055	Bio-Medical Multivariate Statistical Programs (BMD)
000 0056	UNIVAC 1107 Linear Programming Package (LP1107)
000 0057	UNIVAC 1107 PERT/COST, PERT/TIME System (PERT)
000 0058	Inquirer II System for Content Analysis (I/II)
000 0059	Basic Information Retrieval System (BIRS)
000 0060	Matching Factor Solutions (MATCHFS)
000 0061	Matrix Decomposition for Points of View Analysis (MATDEC)
000 0062	Analysis of Change-over Experiments (ZFE-03; ZFE-04)
000 0063	Multivariate Analysis of Variance (MANOVA)
000 0064	Structure-Factor Least-Squares Refinement Program for IBM 7090 (ORFLS-PX)
000 0065	Factor Analysis by Direct "Oblimin" Method (OBLIMIN)
000 0066	Unrestricted Maximum Likelihood Factor Analysis (UMLFA)
000 0067	Rao Constellation and Distance Analysis (DISCRIM2)
000 0068	Nonmetric Multidimensional Scaling (KRUSKAL)
000 0069	Plot (PLOT)
000 0070	Multidimensional Scalogram Analysis (MSA—I)
000 0071	Simulation Package for University Research and Training (SPURT)
000 0072	FORTRAN Program to Assist in the Process of Political Reapportionment (BELOW)
000 0073	Equipercentile Equating Program (SCORMACH)
000 0074	Multiple Scalogram Analysis (MSA)
000 0075	Nonmetric Multidimensional Scaling (KRUSCAL)
000 0076	Algorithms for Analysis of Variance and Covariance of Incomplete Block and Lattice Designs (GAVIAL)
000 0077	An Algorithm for the Optimization of a Quadratic Form Subject to Linear Restraints (ZORILLA)
000 0078	A Regression Model Building System (MOUFLON)
000 0079	TEXT360: A System for Producing Manuals (TEXT360)
000 0080	Synagraphic Mapping Program (SYMAP)
000 0081	FURNIVAL Regression Screen (FURNIVAL)
000 0082	FORTRAN Subroutine Package to Solve Ordinary Differential Equations (NODE)

continued

000 0083	FORTRAN Subroutine Package to Solve Ordinary Differential Equations (DNODE)
000 0084	Bucharest Sort a List into Ascending Order (ABSRT)
000 0085	Bucharest Sort a List into Descending Order (DBSRT)
000 0086	FORTRAN Subroutine to Solve Simultaneous Linear Equations with Complex Coefficeints (CGELG)
000 0087	Minres Method of Factor Analysis (MINRES)
000 0088	Reformat into Either BCD or EBCDIC (EDP C 005 E; EDP C 005 B)
000 0089	Analysis of Contingency Tables (ACT version 1.00)
000 0090	Student Scheduling System (SCHEDULE)
000 0091	Matrix Operations (AES 106)
000 0092	Mineral Identification (MINERAL)
000 0093	Mathematical Programming System (MPS/360)
000 0094	Univariate and Multivariate Analysis of Variance, Covariance and Regression (NYBMUL)
000 0095	University of Maryland Test Scoring Program (Version 4) (UOM 32)
000 0096	Spearman Rank-Order Correlation (UOM 4)
000 0097	Weighted Summing Test Scoring Program (Recoding) (UOM 87)
000 0098	Generalized t Test (UOM 5)
000 0099	Information Retrieval Services
000 0100	University Computing Facility (Marquette)
000 0101	Kiewit Computing Facility (Dartmouth)
000 0102	Remote Job Entry System (RJE)
000 0103	Vogelback Computing Facility
000 0104	MERC Time-Sharing System
000 0105	UCSB On-Line System (OLS)
000 0106	Individual Case Statistics Program (SUMSCRDS)
000 0107	General Multiple Regression (REGAN1)
000 0108	Analysis of Variance Using Between-Subject Designs (SABCA)
000 0109	Numerical Frequency Analysis (NUMFREQ)
000 0110	Simple Multiple Linear Regression (MISREGN)
000 0111	Test-Scoring and Item-Analysis Package (QUICKSCORE/ITEMSTEP/ITEMRS)
000 0112	Page Plotter Using the Line Printer (LPLOT)
000 0113	Smallest Space Analysis (SSA-1)
000 0114	Best Least Squares Polynomial Approximation Subroutine (LSPOL)
000 0115	Stepwise Regression Analysis (STEPREGN)
000 0116	General Program for Multivariate Cross-Classification (NUCROS)
000 0117	Data Survey and Normality Test (NORMSURV)
000 0118	Cross-Cultural Comparison (POLYCOMP)

continued

000 0119	Major Characteristics from Murdock's *Ethnographic Atlas* (ETH—CODE)
000 0120	Random Samples of Cultures (ETH—DGRE)
000 0121	Random Sample of 50 Cultures (ETH—RAND)
000 0122	Demographic Information from Murdock's *Ethnographic Atlas* (ETH—INFO)
000 0123	Comparison of Characteristics from "Ethnographic Atlas" (ETHATLAS)
000 0124	Cultural Comparison (CULTCOMP)
000 0125	Culture Identification (CULTPIK)
000 0126	Finished Characteristics of 400 Cultures (TEXTOR)
000 0127	Finished Characteristics from Textor's *A Cross-Cultural Summary* (TEX—CODE)
000 0128	Fitting Nonlinear Regression Functions (TARSIER)
000 0129	Student's t Test (STUDENTT)
000 0130	Correlation and Multiple Linear Regression (UMST500)
000 0131	Invert Ill-behaved Matrices Using Hotelling and Bodewig's Iterative Technique (HRDMIN)
000 0132	Biomedical Computer Programs (BMD)
000 0133	General Purpose Simulation System (GPSS)
000 0134	Continuous System Modeling Program (CSMP)
000 0135	A Primal-Dual Transportation Algorithm (TRANSPRT)
000 0136	Catholic University Computer Center (Catholic)
000 0137	Chi Square (UMST520)
000 0138	Missing Data Correlation (UMST530)
000 0139	Nonparametric (Rank Order) Statistics (UMST540)
000 0140	FORTRAN Program for Computer Based Serials Holdings Management (UPDATE/HOLD/PUBLSH/LANSUB)
000 0141	Correlation and Orthogonal Factor Analysis (UMST550)
000 0142	Least Square Curve Fitting Using Orthogonal Polynomials (UMST560)
000 0143	Multivariate Analysis of Variance (UMST570)
000 0144	Stepwise Regression (UMST580)
000 0145	Cross Classification (UMST590)
000 0146	Descriptive Statistics (UMST600)
000 0147	Chi Square from Raw Data (UMST620)
000 0148	General Linear Hypothesis for Anova, Unequal Frequencies (UMST610)
000 0149	Single and Simultaneous Equation (TSLS, LISE) Regression Package (UMST630)

KEYWORD INDEX

This Index contains key words from programs listed under the EIN Number Index. This Index should not be considered as exhaustive. It will be added to and updated at regular intervals.

A

- **Administrative**
  - 000 0024 (INFOL)
  - 000 0029 (PERTC)
  - 000 0035 (TSSA)
  - 000 0056 (LP1107)
  - 000 0057 (PERT)
  - 000 0059 (BIRS)
  - 000 0072 (BELOW)
  - 000 0079 (TEXT360)
  - 000 0090 (SCHEDULE)
  - 000 0093 (MPS/360)
  - 000 0135 (TRANSPRT)
- aggregating
  - 000 0039 (SYMAP)
  - 000 0047 (SYMAP)
  - 000 0080 (SYMAP)
- alphabetic
  - 000 0006 (INDEXER)
  - 000 0015 (SCRIPT)
  - 000 0038 (CONTEXT, CONCORD)
  - 000 0041 (DOVACK)
  - 000 0049 (KGIC)
- analog computer
  - 000 0071 (SPURT)
- analysis
  - canonical
    - 000 0055 (BMD)
    - 000 0132 (BMD)
  - change-over experiments
    - 000 0062 (ZFE-03, ZFE-04)
  - cluster
    - 000 0111 (c) (ITEMRS)
  - constellation
    - 000 0067 (DISCRIM2)
  - content
    - 000 0058 (I/II)
- analysis, context
  - 000 0049 (KGIC)
  - 000 0058 (I/II)
  - contingency
    - 000 0055 (BMD)
    - 000 0132 (BMD)
  - correlation
    - 000 0011 (PLOTYY)
    - 000 0012 (BMD29)
    - 000 0014 (NUCORR)
    - 000 0051 (f) (PARCOR)
    - 000 0055 (BMD)
    - 000 0130 (UMST500)
    - 000 0132 (BMD)
    - 000 0141 (UMST550)
  - covariance
    - 000 0005 (MULCVR)
    - 000 0026 (MANOVA)
    - 000 0051 (f) (PARCOR)
    - 000 0055 (BMD)
    - 000 0063 (MANOVA)
    - 000 0076 (GAVIAL)
    - 000 0094 (NYBMUL)
    - 000 0132 (BMD)
    - 000 0143 (UMST570)
  - discriminant
    - 000 0010 (EIDISC)
    - 000 0021 (DISCRIM)
    - 000 0055 (BMD)
    - 000 0063 (MANOVA)
    - 000 0067 (DISCRIM2)
    - 000 0094 (NYBMUL)
    - 000 0132 (BMD)
  - distance
    - 000 0067 (DISCRIM2)
  - factor
    - 000 0002 (FACTOR)
    - 000 0004 (SUZYQ)
    - 000 0027 (MESA1)
    - 000 0035 (TSSA)

continued

- analysis, factor *(cont.)*
  - 000 0051 (i) (FANAL)
  - 000 0051 (p) (ANOVUM)
  - 000 0051 (q) (ANOVES)
  - 000 0051 (r) (AOVRM)
  - 000 0055 (BMD)
  - 000 0065 (OBLIMIN)
  - 000 0066 (UMLFA)
  - 000 0087 (MINRES)
  - 000 0132 (BMD)
  - 000 0141 (UMST550)
- graphic
  - 000 0049 (KGIC)
- Guttman
  - 000 0009 (GUTTSCL)
  - 000 0055 (BMD)
  - 000 0132 (BMD)
- Hoyt
  - 000 0004 (SUZYQ)
- input/output
  - 000 0046 (INOUT)
- item
  - 000 0004 (SUZYQ)
  - 000 0035 (TSSA)
  - 000 0095 (UOM 32)
  - 000 0111 (QUICKSCORE/ ITEMSTEP/ITEMRS)
- linear systems
  - 000 0043 (PALS)
  - 000 0148 (UMST610)
- mathematical
  - 000 0105 (OLS)
- point of view
  - 000 0061 (MATDEC)
- regression
  - 000 0011 (PLOTYY)
  - 000 0012 (BMD29)
  - 000 0026 (MANOVA)
  - 000 0045 (QSASE)
  - 000 0055 (BMD)
  - 000 0063 (MANOVA)
  - 000 0078 (MOUFLON)
  - 000 0094 (NYBMUL)
  - 000 0107 (REGAN1)
  - 000 0115 (STEPREGN)
  - 000 0132 (BMD)
  - 000 0144 (UMST580)
  - 000 0149 (UMST630)
- analysis, scalogram
  - 000 0009 (GUTTSCL)
  - 000 0070 (MSA—I)
  - 000 0074 (MSA)
- score
  - 000 0073 (SCORMACH)
  - 000 0095 (UOM 32)
  - 000 0111 (QUICKSCORE/ ITEMSTEP/ITEMRS)
- smallest space
  - 000 0113 (SSA-1)
- test
  - 000 0004 (SUZYQ)
  - 000 0035 (TSSA)
  - 000 0095 (UOM 32)
  - 000 0111 (QUICKSCORE/ ITEMSTEP/ITEMRS)
- time series
  - 000 0033 (TIMEX)
  - 000 0055 (BMD)
  - 000 0132 (BMD)
- transaction
  - 000 0008 (INDIFF)
- variance
  - 000 0004 (SUZYQ)
  - 000 0005 (MULCVR)
  - 000 0010 (EIDISC)
  - 000 0016 (AID)
  - 000 0026 (MANOVA)
  - 000 0027 (MESA1)
  - 000 0051 (f) (PARCOR)
  - 000 0051 (i) (FANAL)
  - 000 0051 (p) (ANOVUM)
  - 000 0051 (q) (ANOVES)
  - 000 0051 (r) (AOVRM)
  - 000 0055 (BMD)
  - 000 0063 (MANOVA)
  - 000 0076 (GAVIAL)
  - 000 0087 (MINRES)
  - 000 0094 (NYBMUL)
  - 000 0108 (SABCA)
  - 000 0132 (BMD)
  - 000 0143 (UMST570)
  - 000 0148 (UMST610)

- architecture: see Engineering & Technology

continued

art: see Humanities

astronomy: see Physical Sciences

asymptotic regression
000 0055 (BMD)
000 0132 (BMD)

atomic
coordinates
000 0064 (ORFLS-PX)
scattering
000 0064 (ORFLS-PX)
structure
000 0064 (ORFLS-PX)

B

back
solution
000 0067 (DISCRIM2)
000 0130 (UMST500)
substitution
000 0086 (CGELG)

Bartlett's test
000 0042 (CANON)
000 0051 (p) (ANOVUM)
000 0051 (q) (ANOVES)
000 0051 (s) (BARTL)

batch processing
000 0100 (Marquette)
000 0103 (Vogelback)
000 0104 (MERC)
000 0136 (Catholic)

BCD
000 0088 (EDP C 005 E, EDP C 005 B)

Behavioral Sciences
000 0001 (MANNWH)
000 0004 (SUZYQ)
000 0009 (GUTTSCL)
000 0016 (AID)
000 0028 (NUCROS)

Behavioral Sciences *(cont.)*
000 0035 (TSSA)
000 0037 (SAMOS)
000 0039 (SYMAP)
000 0040 (TRAN/PLAN)
000 0041 (DOVACK)
000 0046 (INOUT)
000 0049 (KGIC)
000 0055 (BMD)
000 0058 (I/II)
000 0059 (BIRS)
000 0062 (ZFE-03, ZFE-04)
000 0070 (MSA—I)
000 0072 (BELOW)
000 0074 (MSA)
000 0080 (SYMAP)
000 0093 (MPS/360)
000 0095 (UOM 32)
000 0099
000 0111 (QUICKSCORE/ ITEMSTEP/ITEMRS)
000 0118 (POLYCOMP)
000 0119 (ETH—CODE)
000 0120 (ETH—DGRE)
000 0121 (ETH—RAND)
000 0122 (ETH—INFO)
000 0123 (ETHATLAS)
000 0124 (CULTCOMP)
000 0125 (CULTPIK)
000 0126 (TEXTOR)
000 0127 (TEX—CODE)
000 0132 (BMD)
000 0133 (GPSS)
000 0135 (TRANSPRT)
000 0140 (UPDATE/HOLD/ PUBLSH/LANSUB)

Bessel function
000 0054 (BEEFM)

bibliographic data
000 0034 (TRIAL)

biology: see Life Sciences

biophysics: see Life Sciences

continued

biserial, point
000 0018 (BISR)
000 0035 (TSSA)
000 0095 (UOM 32)

block
  design
    000 0062 (ZFE-03; ZFE-04)
  design, incomplete
    000 0063 (MANOVA)
    000 0076 (GAVIAL)
  diagram
    000 0076 (GAVIAL)
    000 0134 (CSMP)

Bodewig's technique
000 0131 (HRDMIN)

Borgatta error ratio
000 0009 (GUTTSCL)

botany: see Life Sciences

C

canonical correlations
000 0019 (CANON)
000 0042 (CANON)
000 0051 (g) (CANON)
000 0055 (BMD)
000 0063 (MANOVA)
000 0094 (NYBMUL)
000 0132 (BMD)

carry-over
000 0062 (ZFE-03; ZFE-04)

centroid
000 0040 (TRAN/PLAN)
000 0067 (DISCRIM2)

change-over experiments
000 0062 (ZFE-03; ZFE-04)

checking
000 0032 (SEQCHK)

chemistry: see Physical Sciences

chi square
000 0007 (ATTANAL)
000 0008 (INDIFF)
000 0009 (GUTTSCL)
000 0010 (EIDISC)
000 0020 (CHISQR)
000 0028 (NUCROS)
000 0030 (PPRANK)
000 0042 (CANON)
000 0051 (a) (FAWCS)
000 0051 (p) (ANOVUM)
000 0051 (q) (ANOVES)
000 0055 (BMD)
000 0066 (UMLFA)
000 0067 (DISCRIM2)
000 0089 (ACT version 1.00)
000 0116 (NUCROS)
000 0117 (NORMSURV)
000 0137 (UMST520)
000 0139 (UMST540)
000 0145 (UMST590)
000 0147 (UMST620)

Chilton index
000 0009 (GUTTSCL)

cluster analysis
000 0111 (c) (ITEMRS)

coefficients
000 0009 (GUTTSCL)
000 0021 (DISCRIM)
000 0025 (LINPROG)
000 0036 (Z3SLS)
000 0046 (INOUT)
000 0081 (FURNIVAL)
000 0114 (LSPOL)
000 0142 (UMST560)
  biserial
    000 0018 (BISR)
    000 0035 (TSSA)
    000 0095 (UOM 32)
  complex
    000 0050 (FORMAC)
    000 0086 (CGELG)
  concordnace
    000 0030 (PPRANK)
    000 0031 (RKSTAT)
    000 0139 (UMST540)

continued

- coefficients, contingency
  - 000 0020 (CHISQR)
  - 000 0028 (NUCROS)
  - 000 0089 (ACT version 1.00)
  - 000 0116 (NUCROS)
  - 000 0145 (UMST590)
- correlation
  - 000 0007 (ATTANAL)
  - 000 0011 (PLOTYY)
  - 000 0012 (BMD29)
  - 000 0018 (BISR)
  - 000 0027 (MESA1)
  - 000 0031 (RKSTAT)
  - 000 0045 (QSASE)
  - 000 0051 (f) (PARCOR)
  - 000 0051 (l) (KETAU)
  - 000 0051 (m) (SPRHO)
  - 000 0055 (BMD)
  - 000 0065 (OBLIMIN)
  - 000 0071 (SPURT)
  - 000 0075 (KRUSCAL)
  - 000 0087 (MINRES)
  - 000 0095 (UOM 32)
  - 000 0110 (MISREGN)
  - 000 0130 (UMST500)
  - 000 0144 (UMST580)
  - 000 0145 (UMST590)
- discriminant
  - 000 0010 (EIDISC)
  - 000 0021 (DISCRIM)
  - 000 0055 (BMD)
- Kendall rank
  - 000 0028 (NUCROS)
  - 000 0051 (l) (KETAU)
  - 000 0116 (NUCROS)
- linear; colinear
  - 000 0081 (FURNIVAL)
- matrix
  - 000 0043 (PALS)
  - 000 0078 (MOUFLON)
  - 000 0086 (CGELG)
- Pearsonian
  - 000 0008 (INDIFF)
  - 000 0035 (TSSA)
  - 000 0051 (d) (PPMCR)
  - 000 0051 (e) (SIGPP)
- phi
  - 000 0051 (k) (PHICO)

coefficients, phi *(cont.)*

- 000 0095 (UOM 32)
- 000 0111 (a) (QUICKSCORE)
- product—moment
  - 000 0027 (MESA1)
  - 000 0030 (PPRANK)
  - 000 0035 (TSSA)
  - 000 0051 (d) (PPMCR)
  - 000 0051 (e) (SIGPP)
- reproducability
  - 000 0007 (ATTANAL)
  - 000 0009 (GUTTSCL)
- rho
  - 000 0031 (RKSTAT)
  - 000 0051 (m) (SPRHO)
  - 000 0096 (UOM 4)
- Spearman
  - 000 0030 (PPRANK)
  - 000 0031 (RKSTAT)
  - 000 0051 (m) (SPRHO)
  - 000 0096 (UOM 4)
- tau
  - 000 0028 (NUCROS)
  - 000 0051 (l) (KETAU)
  - 000 0116 (NUCROS)
- temperature factor
  - 000 0064 (ORFLX-PX)
- validity
  - 000 0035 (TSSA)

- colinearity
  - 000 0081 (FURNIVAL)

- communality
  - 000 0007 (ATTANAL)
  - 000 0027 (MESA1)
  - 000 0051 (j) (VARMX)
  - 000 0087 (MINRES)
  - 000 0141 (UMST550)

- complex
  - 000 0050 (FORMAC)
  - 000 0086 (CGELG)
  - 000 0105 (OLS)

- components
  - 000 0002 (FACTOR)
  - 000 0027 (MESA1)

continued

- Computer Utility
  - 000 0015 (SCRIPT)
  - 000 0023 (IDSEQN)
  - 000 0032 (SEQCHK)
  - 000 0035 (TSSA)
  - 000 0037 (SAMOS)
  - 000 0039 (SYMAP)
  - 000 0040 (TRAN/PLAN)
  - 000 0047 (SYMAP)
  - 000 0048 (QDGS)
  - 000 0050 (FORMAC)
  - 000 0050 (a) (FMACUT)
  - 000 0052
  - 000 0053 (BEEFDP)
  - 000 0056 (LP1107)
  - 000 0069 (PLOT)
  - 000 0071 (SPURT)
  - 000 0079 (TEXT360)
  - 000 0080 (SYMAP)
  - 000 0084 (ABSRT)
  - 000 0085 (DBSRT)
  - 000 0088 (EDP C 005 B, EDP C 005 E)
  - 000 0091 (AES106)
  - 000 0105 (OLS)
  - 000 0112 (LPLOT)
  - 000 0133 (GPSS)
  - 000 0135 (TRANSPRT)
- concordance
  - 000 0030 (PPRANK)
  - 000 0038 (CONTEXT, CONCORD)
  - coefficients
    - 000 0030 (PPRANK)
    - 000 0031 (RKSTAT)
    - 000 0139 (UMST540)
- confidence level
  - 000 0001 (MANNWH)
- conformant mapping
  - 000 0039 (SYMAP)
  - 000 0047 (SYMAP)
  - 000 0080 (SYMAP)
- consensus
  - 000 0030 (PPRANK)
- consistency
  - index
    - 000 0009 (GUTTSCL)
  - internal
    - 000 0007 (ATTANAL)
    - 000 0032 (SEQCHK)
- constellation analysis
  - 000 0067 (DISCRIM2)
- constraint
  - 000 0025 (LINPROG)
- content analysis
  - 000 0058 (I/II)
- context analysis
  - 000 0049 (KGIC)
  - 000 0058 (I/II)
- contingency
  - 000 0020 (CHISQR)
  - 000 0028 (NUCROS)
  - 000 0051 (a) (FAWCS)
  - 000 0055 (BMD)
  - 000 0089 (ACT version 1.00)
  - 000 0116 (NUCROS)
  - 000 0132 (BMD)
  - 000 0137 (UMST520)
  - 000 0145 (UMST590)
- contour mapping
  - 000 0039 (SYMAP)
  - 000 0047 (SYMAP)
  - 000 0080 (SYMAP)
- contrasts
  - 000 0026 (MANOVA)
  - 000 0055 (BMD)
  - 000 0063 (MANOVA)
  - 000 0132 (BMD)
  - 000 0143 (UMST570)
- Cornell technique
  - 000 0009 (GUTTSCL)
  - 000 0055 (BMD)
- correlation
  - 000 0004 (SUZYQ)
  - 000 0011 (PLOTYY)
  - 000 0012 (BMD29)

continued

correlation *(cont.)*
- 000 0014 (NUCORR)
- 000 0017 (B34T)
- 000 0019 (CANON)
- 000 0027 (MESA1)
- 000 0030 (PPRANK)
- 000 0042 (CANON)
- 000 0045 (QSASE)
- 000 0055 (BMD)
- 000 0065 (OBLIMIN)
- 000 0066 (UMLFA)
- 000 0087 (MINRES)
- 000 0138 (UMST530)
- 000 0141 (UMST550)

coefficients
- 000 0031 (RKSTAT)
- 000 0045 (QSASE)
- 000 0051 (h) (UPREG/DNREG)
- 000 0051 (l) (KETAU)
- 000 0051 (m) (SPRHO)
- 000 0065 (OBLIMIN)
- 000 0071 (SPURT)
- 000 0075 (KRUSCAL)
- 000 0087 (MINRES)
- 000 0096 (UOM 32)
- 000 0110 (MISREGN)
- 000 0111 (QUICKSCORE/ ITEMSTEP/ITEMRS)
- 000 0116 (NUCROS)
- 000 0130 (UMST500)
- 000 0143 (UMST570)
- 000 0145 (UMST590)

canonical
- 000 0019 (CANON)
- 000 0042 (CANON)
- 000 0051 (g) (CANON)
- 000 0055 (BMD)
- 000 0063 (MANOVA)
- 000 0094 (NYBMUL)
- 000 0132 (BMD)

effects
- 000 0026 (MANOVA)

interitem
- 000 0004 (SUZYQ)
- 000 0007 (ATTANAL)
- 000 0035 (TSSA)

correlation, matrix
- 000 0002 (FACTOR)
- 000 0004 (SUZYQ)
- 000 0010 (EIDISC)
- 000 0035 (TSSA)
- 000 0045 (QSASE)
- 000 0067 (DISCRIM2)
- 000 0087 (MINRES)
- 000 0130 (UMST500)
- 000 0141 (UMST550)
- 000 0149 (UMST630)

multiple
- 000 0002 (FACTOR)
- 000 0051 (f) (PARCOR)
- 000 0051 (h) (UPREG/DNREG)

partial
- 000 0051 (f) (PARCOR)
- 000 0051 (h) (UPREG/DNREG)
- 000 0115 (STEPREGN)
- 000 0130 (UMST500)

point-biserial
- 000 0018 (BISR)
- 000 0035 (TSSA)
- 000 0095 (UOM 32)

product-moment
- 000 0027 (MESA1)
- 000 0030 (PPRANK)
- 000 0035 (TSSA)
- 000 0051 (d) (PPMCR)
- 000 0051 (e) (SIGPP)
- 000 0089 (ACT version 1.00)
- 000 0098 (UOM 5)
- 000 0107 (REGAN1)
- 000 0111 (c) (ITEMRS)
- 000 0138 (UMST530)

Spearman
- 000 0139 (UMST540)

covariance
- 000 0005 (MULCVR)
- 000 0010 (EIDISC)
- 000 0026 (MANOVA)
- 000 0051 (f) (PARCOR)
- 000 0055 (BMD)
- 000 0063 (MANOVA)
- 000 0076 (GAVIAL)
- 000 0094 (NYBMUL)

continued

covariance *(cont.)*
- 000 0130 (UMST500)
- 000 0132 (BMD)
- 000 0138 (UMST530)
- 000 0141 (UMST550)
- 000 0143 (UMST570)
- 000 0149 (UMST630)

critical path
- 000 0029 (PERTC)
- 000 0057 (PERT)

crossclassifications
- 000 0020 (CHISQR)
- 000 0028 (NUCROS)
- 000 0051 (a) (FAWCS)
- 000 0055 (BMD)
- 000 0089 (ACT version 1.00)
- 000 0116 (NUCROS)

crosscultural
- 000 0118 (POLYCOMP)
- 000 0119 (ETH–CODE)
- 000 0123 (ETHATLAS)
- 000 0124 (CULTCOMP)
- 000 0127 (TEX–CODE)

crossindexing
- 000 0140 (UPDATE/HOLD/PUBLSH/LANSUB)

crossproducts
- 000 0010 (EIDISC)
- 000 0014 (NUCORR)
- 000 0045 (QSASE)
- 000 0055 (BMD)
- 000 0061 (MATDEC)
- 000 0078 (MOUFLON)
- 000 0098 (UOM 5)
- 000 0130 (UMST500)
- 000 0141 (UMST550)
- 000 0144 (UMST580)
- 000 0149 (UMST630)

crosstabulation
- 000 0020 (CHISQR)
- 000 0028 (NUCROS)
- 000 0051 (a) (FAWCS)
- 000 0055 (BMD)
- 000 0089 (ACT version 1.00)

crosstabulation *(cont.)*
- 000 0116 (NUCROS)
- 000 0145 (UMST590)

cultural, culture
- 000 0118 (POLYCOMP)
- 000 0119 (ETH–CODE)
- 000 0120 (ETH–DGRE)
- 000 0121 (ETH–RAND)
- 000 0122 (ETH–INFO)
- 000 0123 (ETHATLAS)
- 000 0124 (CULTCOMP)
- 000 0125 (CULTPIK)
- 000 0126 (TEXTOR)
- 000 0127 (TEX–CODE)

curve fitting
- 000 0054 (BEEFM)
- 000 0142 (UMST560)

D

decomposition, matrix
- 000 0061 (MATDEC)

degree of freedom
- 000 0007 (ATTANAL)
- 000 0008 (INDIFF)
- 000 0010 (EIDISC)
- 000 0020 (CHISQR)
- 000 0063 (MANOVA)
- 000 0067 (DISCRIM2)
- 000 0087 (MINRES)
- 000 0089 (ACT version 1.00)
- 000 0110 (MISREGN)
- 000 0117 (NORMSURV)
- 000 0139 (UMST540)

deletion
- 000 0012 (BMD29)
- 000 0078 (MOUFLON)

demographic
- 000 0072 (BELOW)
- 000 0122 (ETH–INFO)

density, population
- 000 0039 (SYMAP)

continued

- density, population *(cont.)*
  - 000 0047 (SYMAP)
  - 000 0072 (BELOW)
  - 000 0080 (SYMAP)
- design
  - block
    - 000 0062 (ZFE-03, ZFE-04)
    - 000 0076 (GAVIAL)
    - 000 0134 (CSMP)
  - crossed
    - 000 0094 (NYBMUL)
  - incomplete
    - 000 0063 (MANOVA)
    - 000 0076 (GAVIAL)
  - lattice
    - 000 0076 (GAVIAL)
  - nested
    - 000 0094 (NYBMUL)
- detection
  - 000 0016 (AID)
- determinant
  - 000 0022 (EIGSYS)
  - 000 0054 (BEEFM)
  - 000 0055 (BMD)
- determination
  - 000 0081 (FURNIVAL)
- dictation
  - 000 0041 (DOVACK)
- dictionary
  - 000 0038 (CONTEXT, CONCORD)
  - 000 0041 (DOVACK)
  - 000 0058 (I/II)
  - 000 0079 (TEXT360)
- differences, group
  - 000 0010 (EIDISC)
- differential equations
  - 000 0043 (PALS)
  - 000 0054 (BEEFM)
  - 000 0082 (NODE)
  - 000 0083 (DNODE)
  - 000 0134 (CSMP)
- dimensions
  - 000 0001 (MANNWH)
  - 000 0028 (NUCROS)
- dimensions *(cont.)*
  - 000 0061 (MATDEC)
  - 000 0068 (KRUSKAL)
  - 000 0070 (MSA–I)
  - 000 0075 (KRUSCAL)
  - 000 0113 (SSA-1)
- discriminant analysis; functions
  - 000 0010 (EIDISC)
  - 000 0021 (DISCRIM)
  - 000 0055 (BMD)
  - 000 0063 (MANOVA)
  - 000 0067 (DISCRIM2)
  - 000 0094 (NYBMUL)
  - 000 0132 (BMD)
- discrimination index
  - 000 0095 (UOM 32)
  - 000 0111 (a) (QUICKSCORE)
  - 000 0111 (b) (ITEMSTEP)
- dispersion
  - 000 0064 (ORFLS-PX)
- distance
  - analysis
    - 000 0067 (DISCRIM2)
  - statistic
    - 000 0010 (EIDISC)
- distribution
  - Erlang
    - 000 0071 (SPURT)
  - frequency
    - 000 0004 (SUZYQ)
    - 000 0009 (GUTTSCL)
    - 000 0051 (a) (FAWCS)
    - 000 0071 (SPURT)
    - 000 0087 (MINRES)
    - 000 0089 (ACT version 1.00)
    - 000 0117 (NORMSURV)
    - 000 0145 (UMST590)
    - 000 0147 (UMST620)
  - Poisson
    - 000 0071 (SPURT)
  - uniform
    - 000 0054 (BEEFM)
    - 000 0071 (SPURT)
- district, political
  - 000 0072 (BELOW)

continued

D, Somers
  000 0028 (NUCROS)
  000 0116 (NUCROS)

D^2, Mahalanobis
  000 0010 (EIDISC)

dual algorithm
  000 0056 (LP1107)
  000 0135 (TRANSPRT)

Durbin-Watson statistic
  000 0012 (BMD29)
  000 0130 (UMST500)
  000 0149 (UMST630)

E

Earth Sciences
  000 0039 (SYMAP)
  000 0047 (SYMAP)
  000 0080 (SYMAP)
  000 0092 (MINERAL)

econometric
  000 0036 (Z3SLS)
  000 0046 (INOUT)

edit(ing)
  000 0034 (TRIAL)

education: see **Behavioral Sciences**

eigenvalue
  000 0010 (EIDISC)
  000 0022 (EIGSYS)
  000 0027 (MESA1)
  000 0043 (PALS)
  000 0051 (g) (CANON)
  000 0051 (i) (FANAL)
  000 0061 (MATDEC)
  000 0141 (UMST550)

eigenvector
  000 0010 (EIDISC)
  000 0022 (EIGSYS)
  000 0043 (PALS)

eigenvector *(cont.)*
  000 0051 (g) (CANON)
  000 0051 (i) (FANAL)
  000 0061 (MATDEC)

Engineering & Technology
  000 0027 (MESA1)
  000 0039 (SYMAP)
  000 0043 (PALS)
  000 0047 (SYMAP)
  000 0054 (BEEFM)
  000 0071 (SPURT)
  000 0080 (SYMAP)
  000 0134 (CSMP)

equating
  000 0073 (SCORMACH)

equipercentile
  000 0073 (SCORMACH)

Erlang distribution
  000 0071 (SPURT)

error
  Borgatta error ratio
    000 0009 (GUTTSCL)
  counting
    000 0009 (GUTTSCL)
  standard
    000 0012 (BMD29)
    000 0027 (MESA1)
    000 0033 (TIMEX)
    000 0035 (TSSA)
    000 0036 (Z3SLS)
    000 0045 (QSASE)
    000 0051 (b) (STSUM)
    000 0051 (c) (TTEST)
    000 0055 (BMD)
    000 0071 (SPURT)
    000 0078 (MOUFLON)
    000 0107 (REGAN1)
    000 0109 (NUMFREQ)
    000 0110 (MISREGN)
    000 0129 (STUDENTT)
    000 0130 (UMST500)
    000 0141 (UMST550)
    000 0142 (UMST560)
    000 0149 (UMST630)

continued

ethnography
- 000 0118 (POLYCOMP)
- 000 0119 (ETH–CODE)
- 000 0120 (ETH–DGRE)
- 000 0121 (ETH–RAND)
- 000 0122 (ETH–INFO)
- 000 0123 (ETHATLAS)
- 000 0124 (CULTCOMP)
- 000 0125 (CULTPIK)
- 000 0126 (TEXTOR)
- 000 0127 (TEX–CODE)

Euclidean
- 000 0070 (MSA–I)
- 000 0113 (SSA–1)

evaluation
- 000 0041 (DOVACK)

exclusion
- 000 0044 (PROFILE)
- 000 0081 (FURNIVAL)

F

facility
- 000 0100 (Marquette)
- 000 0101 (Dartmouth)
- 000 0102 (RJE)
- 000 0103
- 000 0104
- 000 0105 (OLS)

factor
- 000 0002 (FACTOR)
- 000 0004 (SUZYQ)
- 000 0007 (ATTANAL)
- 000 0027 (MESA1)
- 000 0060 (MATCHFS)
- 000 0063 (MANOVA)
- 000 0066 (UMLFA)
- analysis
  - 000 0004 (SUZYQ)
  - 000 0035 (TSSA)
  - 000 0051 (i) (FANAL)
  - 000 0051 (p) (ANOVUM)
  - 000 0051 (q) (ANOVES)

factor, analysis *(cont.)*
  - 000 0051 (r) (AOVRM)
  - 000 0055 (BMD)
  - 000 0065 (OBLIMIN)
  - 000 0087 (MINRES)
  - 000 0132 (BMD)
  - 000 0141 (UMST550)
- loading
  - 000 0002 (FACTOR)
  - 000 0007 (ATTANAL)
  - 000 0027 (MESA1)
  - 000 0051 (j) (VARMX)
  - 000 0087 (MINRES)
- solution, matching
  - 000 0060 MATCHFS)
- structure
  - 000 0064 (ORFLS-PX)

file
- creation & maintenance
  - 000 0034 (TRIAL)
  - 000 0059 (BIRS)
  - 000 0140 (UPDATE/HOLD/ PUBLSH/LANSUB)
- referencing
  - 000 0024 (INFOL)
  - 000 0140 (UPDATE/HOLD/ PUBLSH/LANSUB)

Fisher
- 000 0035 (TSSA)
- 000 0111 (c) (ITEMRS)

fit, goodness of
- 000 0008 (INDIFF)
- 000 0009 (GUTTSCL)
- 000 0066 (UMLFA)

flow, transaction
- 000 0008 (INDIFF)

font
- 000 0015 (SCRIPT)

FORMAC
- 000 0050 (FORMAC)
- 000 0050 (a) (FMACUT)

frequency
- 000 0009 (GUTTSCL)

continued

frequency *(cont.)*
000 0041 (DOVACK)
000 0044 (PROFILE)
000 0051 (c) (TTEST)
000 0071 (SPURT)
000 0094 (NYBMUL)
000 0095 (UOM 32)
000 0109 (NUMFREQ)
000 0137 (UMST520)
000 0146 (UMST600)
distribution
000 0004 (SUZYQ)
000 0009 (GUTTSCL)
000 0051 (a) (FAWCS)
000 0071 (SPURT)
000 0087 (MINRES)
000 0089 (ACT version 1.00)
000 0117 (NORMSURV)
000 0145 (UMST590)
000 0147 (UMST620)
load
000 0043 (PALS)
response
000 0009 (GUTTSCL)
000 0055 (BMD)
tables
000 0137 (UMST520)

Fresnel
000 0054 (BEEFM)

F ratio; test; value
000 0003 (STDNTF2)
000 0010 (EIDISC)
000 0012 (BMD29)
000 0033 (TIMEX)
000 0045 (QSASE)
000 0051 (h) (UPREG/DNREG)
000 0055 (BMD)
000 0078 (MOUFLON)
000 0098 (UOM 5)
000 0107 (REGAN1)
000 0110 (MISREGN)
000 0115 (STEPREGN)
000 0129 (STUDENTT)
000 0144 (UMST580)
000 0148 (UMST610)
000 0149 (UMST630)

function
000 0050 (FORMAC)
000 0050 (a) (FMACUT)
000 0054 (BEEFM)
Bessel
000 0054 (BEEFM)
gamma
000 0054 (BEEFM)
regression
000 0128 (TARSIER)

G

gamma
function
000 0054 (BEEFM)
Goodman-Kurskal
000 0028 (NUCROS)
000 0089 (ACT version 1.00)
000 0116 (NUCROS)
000 0145 (UMST590)

Gauss
000 0054 (BEEFM)
000 0086 (CGELG)
000 0087 (MINRES)

Gauss—Newton method
000 0128 (TARSIER)

General Facility Access
000 0100 (Marquette)
000 0101 (Dartmouth)
000 0102 (RJE)
000 0103
000 0104
000 0105 (OLS)
000 0136 (Catholic)

geography: see Earth Sciences

geology: see Earth Sciences

Goodman
000 0008 (INDIFF)
000 0009 (GUTTSCL)

Goodman—Kruskal gamma
000 0028 (NUCROS)

continued

Goodman–Kruskal gamma *(cont.)*
  000 0089 (ACT version 1.00)
  000 0116 (NUCROS)
  000 0145 (UMST590)

goodness of fit
  000 0008 (INDIFF)
  000 0009 (GUTTSCL)
  000 0066 (UMLFA)
  000 0075 (KRUSCAL)

gradient
  000 0075 (KRUSCAL)

grapheme
  000 0049 (KGIC)

graphic analysis
  000 0049 (KGIC)

graphics, graphs
  000 0015 (SCRIPT)
  000 0039 (SYMAP)
  000 0047 (SYMAP)
  000 0048 (QDGS)
  000 0052
  000 0069 (PLOT)
  000 0071 (SPURT)
  000 0080 (SYMAP)
  000 0105 (OLS)
  000 0112 (LPLOT)
  000 0142 (UMST560)
  000 0149 (UMST630)

Green
  consistency index
    000 0009 (GUTTSCL)
  response patterns
    000 0009 (GUTTSCL)

group
  differences
    000 0010 (EIDISC)
  interactions
    000 0075 (KRUSCAL)
  membership
    000 0010 (EIDISC)
  variability
    000 0067 (DISCRIM2)

Guilford index
  000 0035 (TSSA)

Gullicksen index
  000 0035 (TSSA)

Guttman
  analysis
    000 0009 (GUTTSCL)
    000 0055 (BMD)
    000 0132 (BMD)
  unidimensionality requirement
    000 0007 (ATTANAL)

H

Hilbert matrix
  000 0131 (HRDMIN)

histogram
  000 0010 (EIDISC)
  000 0041 (DOVACK)
  000 0071 (SPURT)
  000 0095 (UOM 32)
  000 0111 (QUICKSCORE/ ITEMSTEP/ITEMRS)

history: see Humanities

H, Kruskal–Wallis
  000 0031 (RKSTAT)
  000 0051 (o) (KRWAL)
  000 0089 (ACT version 1.00)

Hocking–Leslie statistic
  000 0078 (MOUFLON)

holdings
  000 0140 (UPDATE/HOLD PUBLSH/LANSUB)

homogeneity
  index
    000 0009 (GUTTSCL)
  test, Bartlett's
    000 0051 (p) (ANOVUM)
    000 0051 (q) (ANOVES)
    000 0051 (s) (BARTL)

Hotelling
  000 0005 (MULCVR)
  000 0019 (CANON)

continued

Hotelling *(cont.)*
000 0131 (HRDMIN)
000 0144 (UMST580)

Hoyt, analysis of variance
000 0004 (SUZYQ)

Humanities
000 0038 (CONTEXT, CONCORD)
000 0049 (KGIC)
000 0058 (I/II)

I

identification
000 0023 (IDSEQN)
mineral
000 0092 (MINERAL)

index, indexing
000 0006 (INDEXER)
000 0034 (TRIAL)
000 0140 (UPDATE/HOLD PUBLSH/LANSUB)
Chilton
000 0009 (GUTTSCL)
consistency
000 0009 (GUTTSCL)
difficulty
000 0095 (UOM 32)
discrimination
000 0095 (UOM 32)
000 0111 (a) (QUICKSCORE)
000 0111 (b) (ITEMSTEP)
Green
000 0009 (GUTTSCL)
Guilford, Gullicksen
000 0035 (TSSA)
homogeneity, Loevinger
000 0009 (GUTTSCL)
Raju
000 0009 (GUTTSCL)
reliability
000 0035 (TSSA)
Kuder-Richardson
000 0009 (GUTTSCL)

index, validity
000 0035 (TSSA)
W
000 0111 (b) (ITEMSTEP)
000 0111 (c) (ITEMRS)

information
processing
000 0034 (TRIAL)
000 0053 (BEEFDP)
000 0058 (I/II)
000 0140 (UPDATE/HOLD/ PUBLSH/LANSUB)
retrieval
000 0024 (INFOL)
000 0034 (TRIAL)
000 0059 (BIRS)
000 0099
000 0140 (UPDATE/HOLD/ PUBLSH/LANSUB)

information sciences: see Library & Information Sciences

instruction
000 0037 (SAMOS)
000 0041 (DOVACK)

integration
000 0054 (BEEFM)

interaction
000 0008 (INDIFF)
000 0016 (AID)
000 0143 (UMST570)

intercorrelation
000 0002 (FACTOR)
000 0007 (ATTANAL)
000 0042 (CANON)
000 0067 (DISCRIM2)

interdependency
000 0046 (INOUT)

interitem correlation
000 0004 (SUZYQ)
000 0007 (ATTANAL)
000 0035 (TSSA)

continued

internal consistency
000 0007 (ATTANAL)
000 0032 (SEQCHK)

interpolation
000 0054 (BEEFM)

interpreter
000 0050 (FORMAC)
000 0050 (a) (FMACUT)

interrelation
000 0016 (AID)
000 0042 (CANON)

interrogation
000 0024 (INFOL)

item analysis
000 0004 (SUZYQ)
000 0035 (TSSA)
000 0095 (UOM 32)
000 0111 (QUICKSCORE/ITEMSTEP/ITEMRS)

iteration, inverse
000 0043 (PALS)

inverse matrix
000 0130 (UMST500)
000 0131 (HRDMIN)

J

Jacobian
000 0027 (MESA1)
000 0051 (g) (CANON)

justification line
000 0079 (TEXT360)

K

Kaiser: see varimax

Kendall tau
000 0028 (NUCROS)
000 0051 (1) (KETAU)

Kendall tau *(cont.)*
000 0116 (NUCROS)
000 0139 (UMST540)
000 0145 (UMST590)
W
000 0030 (PPRANK)
000 0031 (RKSTAT)
000 0139 (UMST540)

key word(s)
000 0034 (TRIAL)

—Kruskal: see Goodman

Kruskal scaling
000 0068 (KRUSKAL)
000 0075 (KRUSCAL)

Kruskal-Wallis H
000 0031 (RKSTAT)
000 0051 (o) (KRWAL)
000 0089 (ACT version 1.00)
000 0139 (UMST540)

Kruder
preference record
000 0035 (TSSA)
—Richardson reliability index
000 0007 (ATTANAL)
000 0009 (GUTTSCL)
000 0035 (TSSA)
000 0095 (UOM 32)

kurtosis
000 0035 (TSSA)
000 0051 (b) (STSUM)

Kutta: see Runge

L

LaGrange
000 0054 (BEEFM)

lambda, Wilk
000 0010 (EIDISC)

language: see Computer
Utility
FORMAC

continued

Latin squares
000 0062 (ZFE-03, ZFE-04)

lattice designs
000 0076 (GAVIAL)

Lawley
chi square
000 0066 (UMLFA)
test of significance
000 0007 (ATTANAL)

least squares
000 0036 (Z3SLS)
000 0045 (QSASE)
000 0060 (MATCHFS)
000 0062 (ZFE-03, ZFE-04)
000 0094 (NYBMUL)
000 0142 (UMST560)
000 0148 (UMST610)
000 0149 (UMST630)

legislature
000 0072 (BELOW)

Leontief
000 0046 (INOUT)

Leslie: see Hocking

level of confidence
000 0001 (MANNWH)

Library & Information Sciences
000 0006 (INDEXER)
000 0024 (INFOL)
000 0034 (TRIAL)
000 0049 (KGIC)
000 0053 (BEEFDP)
000 0058 (I/II)
000 0059 (BIRS)
000 0079 (TEXT360)
000 0099
000 0140 (UPDATE/HOLD/
PUBLSH/LANSUB)

Life Sciences
000 0016 (AID)
000 0055 (BMD)
000 0062 (ZFE-03, ZFE-04)
000 0099

Life Sciences *(cont.)*
000 0132 (BMD)
000 0140 (UPDATE/HOLD/
PUBLSH/LANSUB)

likelihood, maximum
000 0014 (NUCORR)
000 0045 (QSASE)
000 0055 (BMD)
000 0066 (UMLFA)
000 0149 (UMAT630)

line justification
000 0079 (TEXT360)

linear
hypothesis
000 0148 (UMST610)
programming
000 0025 (LINPROG)
000 0043 (PALS)
000 0056 (LP1107)
000 0093 (MPS/360)
regression
000 0055 (BMD)
000 0130 (UMST500)
000 0132 (BMD)
000 0144 (UMST580)

linearity coefficient
000 0081 (FURNIVAL)

list(ing)
000 0023 (IDSEQN)
000 0034 (TRIAL)
000 0038 (CONTEXT, CONCORD)
000 0073 (SCORMACH)
000 0084 (ABSRT)
000 0085 (DBSRT)

literature, scientific
000 0099

load frequency
000 0043 (PALS)

loading, factor
000 0002 (FACTOR)
000 0027 (MESA1)
000 0051 (j) (VARMX)
000 0087 (MINRES)

continued

loci, root
000 0043 (PALS)

Loevinger
homogeneity index
000 0009 (GUTTSCL)
response patterns
000 0009 (GUTTSCL)

M

Mahalanobis
000 0010 (EIDISC)

management, holdings
000 00140 (UPDATE/HOLD/PUBLSH/LANSUB)

Mann–Whitney U test
000 0001 (MANNWH)
000 0031 (RKSTAT)
000 0051 (n) (MNWHT)

manuals
000 0079 (TEXT360)

map, mapping
000 0048 (QDGS)
000 0070 (MSA–I)
conformant
000 0039 (SYMAP)
000 0047 (SYMAP)
000 0080 (SYMAP)
contour
000 0039 (SYMAP)
000 0047 (SYMAP)
000 0080 (SYMAP)
proximal
000 0039 (SYMAP)
000 0047 (SYMAP)
000 0080 (SYMAP)

marginals
000 0009 (GUTTSCL)

matching factor solutions
000 0060 (MARCHFS)

Mathematics
000 0022 (EIGSYS)
000 0025 (LINPROG)
000 0027 (MESA1)
000 0043 (PALS)
000 0045 (QSASE)
000 0050 (FORMAC)
000 0050 (a) (FMACUT)
000 0054 (BEEFM)
000 0060 (MATCHFS)
000 0061 (MATDEC)
000 0065 (OBLIMIN)
000 0067 (DISCRIM2)
000 0068 (KRUSKAL)
000 0071 (SPURT)
000 0075 (KRUSCAL)
000 0082 (NODE)
000 0083 (DNODE)
000 0086 (CGELG)
000 0087 (MINRES)
000 0091 (AES106)
000 0105 (OLS)
000 0113 (SSA-1)
000 0114 (LSPOL)
000 0131 (HRDMIN)
000 0133 (GPSS)
000 0134 (CSMP)

matrix
000 0008 (INDIFF)
000 0022 (EIGSYS)
000 0027 (MESA1)
000 0042 (CANON)
000 0043 (PALS)
000 0054 (BEEFM)
000 0055 (BMD)
000 0061 (MATDEC)
000 0071 (SPURT)
000 0075 (KRUSCAL)
000 0077 (ZORILLA)
000 0078 (MOUFLON)
000 0086 (CGELG)
000 0087 (MINRES)
000 0091 (AES106)
000 0096 (UOM 4)
000 0107 (REGAN1)

continued

- matrix *(cont.)*
  - 000 0111 (QUICKSCORE/ITEMSTEP/ITEMRS)
  - 000 0130 (UMST500)
  - 000 0131 (HRDMIN)
  - 000 0138 (UMST530)
  - 000 0149 (UMST630)
  - correlation
    - 000 0002 (FACTOR)
    - 000 0004 (SUZYQ)
    - 000 0010 (EIDISC)
    - 000 0035 (TSSA)
    - 000 0045 (QSASE)
    - 000 0067 (DISCRIM2)
    - 000 0075 (KRUSCAL)
    - 000 0087 (MINRES)
    - 000 0130 (UMST500)
    - 000 0141 (UMST550)
    - 000 0144 (UMST580)
  - decomposition
    - 000 0061 (MATDEC)
  - dissimilarities
    - 000 0075 (KRUSCAL)
  - intercorrelation
    - 000 0067 (DISCRIM2)
  - inversion
    - 000 0131 (HRDMIN)
  - orthogonal factor
    - 000 0141 (UMST550)
  - residuals
    - 000 0007 (ATTANAL)
    - 000 0012 (BMD29)
    - 000 0078 (MOUFLON)
  - target
    - 000 0060 (MATCHFS)
  - variance
    - 000 0143 (UMST570)
    - 000 0149 (UMST630)
- maximum likelihood
  - 000 0014 (NUCORR)
  - 000 0045 (QSASE)
  - 000 0055 (BMD)
  - 000 0066 (UMLFA)
  - 000 0132 (BMD)
- means
  - 000 0002 (FACTOR)
  - 000 0007 (ATTANAL)
  - 000 0008 (INDIFF)
  - 000 0010 (EIDISC)
- means *(cont.)*
  - 000 0014 (NUCORR)
  - 000 0018 (BISR)
  - 000 0026 (MANOVA)
  - 000 0027 (MESA1)
  - 000 0035 (TSSA)
  - 000 0044 (PROFILE)
  - 000 0045 (QSASE)
  - 000 0051 (b) (STSUM)
  - 000 0051 (c) (TTEST)
  - 000 0051 (d) (PPMCR)
  - 000 0062 (ZFE-03, ZFE-04)
  - 000 0067 (DISCRIM2)
  - 000 0071 (SPURT)
  - 000 0078 (MOUFLON)
  - 000 0089 (ACT version 1.00)
  - 000 0094 (NYBMUL)
  - 000 0095 (UOM 32)
  - 000 0098 (UOM 5)
  - 000 0106 (SUMSCRDS)
  - 000 0107 (REGAN1)
  - 000 0108 (SABCA)
  - 000 0109 (NUMFREQ)
  - 000 0110 (MISREGN)
  - 000 0111 (a) (QUICKSCORE)
  - 000 0117 (NORMSURV)
  - 000 0129 (STUDENTT)
  - 000 0130 (UMST500)
  - 000 0138 (UMST530)
  - 000 0141 (UMST550)
  - 000 0142 (UMST560)
  - 000 0146 (UMST600)
  - 000 0148 (UMST610)
  - 000 0149 (UMST630)
- measurement: see Statistics & Measurement
- median
  - 000 0001 (MANNWH)
  - 000 0008 (INDIFF)
  - 000 0044 (PROFILE)
  - 000 0117 (NORMSURV)
- medicine: see Life Sciences
- membership probabilities
  - 000 0010 (EIDISC)
- Menzel
  - 000 0009 (GUTTSCL)

continued

merge
000 0041 (DOVACK)
000 0084 (ABSRT)
000 0085 (DBSRT)

mineralogy: see Earth Sciences

minerals
000 0092 (MINERAL)

minimizing
000 0025 (LINPROG)
000 0029 (PERTC)
000 0056 (LP1107)
000 0057 (PERT)
000 0065 (OBLIMIN)
000 0077 (ZORILLA)
000 0078 (MOUFLON)
000 0087 (MINRES)
000 0093 (MPS/360)
000 0135 (TRANSPRT)

modelling
000 0071 (SPURT)
000 0078 (MOUFLON)
000 0133 (GPSS)
000 0134 (CSMP)
000 0135 (TRANSPRT)

moment
000 0027 (MESA1)
000 0051 (b) (STSUM)
000 0146 (UMST600)
000 0149 (UMST630)
  gross
000 0081 (FURNIVAL)
  population
000 0072 (BELOW)
  product–
000 0027 (MESA1)
000 0030 (PPRANK)
000 0035 (TSSA)
000 0051 (d) (PPMCR)
000 0051 (e) (SIGPP)
000 0089 (ACT version 1.00)
000 0098 (UOM 5)
000 0107 (REGAN1)
000 0111 (c) (ITEMRS)
000 0138 (UMST530)

Mood test
000 0033 (TIMEX)

multidimensional
  scaling
000 0061 (MATDEC)
000 0068 (KRUSKAL)
000 0075 (KRUSCAL)
  scalogram analysis
000 0070 (MSA–I)

multiple
  correlation
000 0002 (FACTOR)
000 0051 (f) (PARCOR)
000 0051 (h) (UPREG/DNREG)
  scalogram analysis
000 0074 (MSA)

music: see Humanities

N

Nagel
000 0072 (BELOW)

nearest neighbor
000 0039 (SYMAP)
000 0047 (SYMAP)

network
000 0029 (PERTC)
000 0040 (TRAN/PLAN)
000 0043 (PALS)
000 0057 (PERT)

neutron
000 0064 (ORFLS-PX)

nonlinear functions
000 0128 (TARSIER)
000 0134 (CSMP)

nonmetric
000 0068 (KRUSKAL)
000 0075 (KRUSCAL)
000 0113 (SSA-1)

continued

nonparametric
    000 0001 (MANWH)
    000 0030 (PPRANK)
    000 0139 (UMST540)
    000 0145 (UMST590)

O

O'Niell
    000 0038 (CONTEXT, CONCORD)

on-line
    000 0101
    000 0102 (RJE)
    000 0103
    000 0104
    000 0105 (OLS)

Operations Research
    000 0008 (INDIFF)
    000 0025 (LINPROG)
    000 0029 (PERTC)
    000 0039 (SYMAP)
    000 0040 (TRAN/PLAN)
    000 0043 (PALS)
    000 0045 (QSASE)
    000 0046 (INOUT)
    000 0047 (SYMAP)
    000 0056 (LP1107)
    000 0057 (PERT)
    000 0071 (SPURT)
    000 0077 (ZORILLA)
    000 0078 (MOUFLON)
    000 0080 (SYMAP)
    000 0086 (CGELG)
    000 0093 (MPS/360)
    000 0133 (GPSS)
    000 0134 (CSMP)
    000 0135 (TRANSPRT)

operator, Boolean
    000 0034 (TRIAL)

orthogonal
    000 0026 (MANOVA)
    000 0060 (MATCHFS)
    000 0143 (UMST570)

orthogonal factors
    000 0141 (UMST550)
  polynomials
    000 0142 (UMST560)

order(ing)
    000 0032 (SEQCHK)
    000 0070 (MSA–I)
    000 0084 (ABSRT)
    000 0085 (DBSRT)
    000 0113 (SSA-1)
    000 0139 (UMST540)
  ascending
    000 0023 (IDSEQN)
    000 0084 (ABSRT)
  descending
    000 0085 (DBSRT)

P

parametric
    000 0055 (BMD)
    000 0132 (BMD)

partial
  correlation
    000 0051 (f) (PARCOR)
    000 0051 (h) (UPREG/DNREG)
    000 0115 (STEPREGN)
    000 0130 (UMST500)
  regression
    000 0149 (UMST630)

partitioning
    000 0042 (CANON)
    000 0129 (STUDENTT)

Pearson
    000 0107 (REGAN1)
  coefficient
    000 0008 (INDIFF)
    000 0035 (TSSA)
    000 0051 (d) (PPMCR)
    000 0051 (e) (SIGPP)
    000 0138 (UMST530)

percentages, row and column
    000 0027 (MESA1)

continued

percentages, row and col. *(cont.)*
000 0028 (NUCROS)
000 0116 (NUCROS)

percentile
000 0073 (SCORMACH)

phi coefficient
000 0051 (k) (PHICO)
000 0095 (UOM 32)
000 0111 (a) (QUICKSCORE)

phonetics, phoneme
000 0049 (KGIC)

Physical Sciences
000 0064 (ORFLS-PX)
000 0099

physics: see Physical Sciences

physiology: see Life Sciences

pivoting
000 0086 (CGELG)

planning
student
000 0090 (SCHEDULE)
time and cost
000 0029 (PERTC)
000 0057 (PERT)
000 0135 (TRANSPRT)
transportation
000 0040 (TRAN/PLAN)
000 0135 (TRANSPRT)

PL/I
000 0050 (FORMAC)
000 0050 (a) (FMACUT)

plots
000 0010 (EIDISC)
000 0011 (PLOTYY)
000 0048 (QDGS)
000 0052
000 0069 (PLOT)
000 0070 (MSA—I)
000 0071 (SPURT)
000 0075 (KRUSCAL)
000 0112 (LPLOT)
000 0145 (UMST590)

point biserial
000 0018 (BISR)
000 0035 (TSSA)
000 0095 (UOM 32)

points of view analysis
000 0061 (MATDEC)

Poisson distribution
000 0071 (SPURT)

political distribution
000 0072 (BELOW)

polynomials
000 0026 (MANOVA)
000 0036 (Z3SLS)
000 0054 (BEEFM)
000 0063 (MANOVA)
000 0114 (LSPOL)
000 0132 (BMD)
000 0142 (UMST560)

population
density
000 0039 (SYMAP)
000 0047 (SYMAP)
000 0072 (BELOW)
000 0080 (SYMAP)
moment
000 0072 (BELOW)

power
000 0043 (PALS)
000 0114 (LSPOL)

predictor(s)
000 0016 (AID)
000 0042 (CANON)
000 0081 (FURNIVAL)
000 0082 (NODE)
000 0083 (DNODE)

preference record, Kuder
000 0035 (TSSA)

principal components
000 0002 (FACTOR)

print(ing)
000 0006 (INDEXER)
000 0079 (TEXT360)
000 0112 (LPLOT)
000 0140 (UPDATE/HOLD/
PUBLSH/LANSUB)

continued

probabilities
000 0009 (GUTTSCL)
000 0129 (STUDENTT)
000 0144 (UMST580)
group membership
000 0010 (EIDISC)

processing
information
000 0034 (TRIAL)
000 0053 (BEEFDP)
000 0058 (I/II)
text
000 0079 (TEXT360)

product–moment
000 0027 (MESA1)
000 0030 (PPRANK)
000 0035 (TSSA)
000 0051 (d) (PPMCR)
000 0051 (e) (SIGPP)
000 0089 (ACT version 1.00)
000 0098 (UOM 5)
000 0107 (REGAN1)
000 0111 (c) (ITEMRS)
000 0138 (UMST530)

programming
000 0037 (SAMOS)
linear
000 0025 (LINPROG)
000 0043 (PALS)
000 0056 (LP1107)
000 0093 (MPS/360)
mathematical
000 0093 (MPS/360)

proximal mapping
000 0039 (SYMAP)
000 0047 (SYMAP)
000 0080 (SYMAP)

psychology: see Behavioral Sciences

psychometric scores
000 0007 (ATTANAL)
000 0009 (GUTTSCL)

Q

quartimax matrix
000 0141 (UMST550)

questionnaire
000 0004 (SUZYQ)

queueing
000 0133 (GPSS)

QR transform
000 0043 (PALS)
000 0081 (FURNIVAL)
000 0086 (CGELG)

R

Raju index
000 0009 (GUTTSCL)

rank
000 0028 (NUCROS)
000 0030 (PPRANK)
000 0031 (RKSTAT)
000 0051 (l) (KETAU)
000 0051 (m) (SPRHO)
000 0073 (SCORMACH)
000 0096 (UOM 4)
000 0113 (SSA-1)
000 0116 (NUCROS)
000 0139 (UMST540)

Rao
000 0010 (EIDISC)
000 0021 (DISCRIM)
000 0067 (DISCRIM2)

rating
000 0030 (PPRANK)
000 0096 (UOM 4)

Rayleigh
000 0022 (EIGSYS)
000 0043 (PALS)

reading
beginning, remedial
000 0041 (DOVACK)

continued

reanalysis
000 0026 (MANOVA)
000 0063 (MANOVA)

reapportionment, political
000 0072 (BELOW)

recoding
000 0088 (EDP C 005 B, EDP C 005 E)
000 0097 (UOM 87)

recognition tests
000 0041 (DOVACK)

redundancy measures
000 0042 (CANON)

reformat
000 0088 (EDP C 005 B, EDP C 005 E)

regression
000 0017 (B34T)
000 0033 (TIMEX)
000 0036 (Z3SLS)
000 0045 (QSASE)
000 0051 (h) (UPREG/DNREG)
000 0055 (BMD)
000 0078 (MOUFLON)
000 0081 (FURNIVAL)
000 0110 (MISREGN)
analysis
000 0011 (PLOTYY)
000 0012 (BMD29)
000 0026 (MANOVA)
000 0045 (QSASE)
000 0055 (BMD)
000 0063 (MANOVA)
000 0078 (MOUFLON)
000 0094 (NYBMUL)
000 0107 (REGAN1)
000 0115 (STEPREGN)
000 0132 (BMD)
000 0144 (UMST580)
000 0149 (UMST630)
asymptotic
000 0055 (BMD)
000 0132 (BMD)
functions, nonlinear
000 0128 (TARSIER)

regression, linear
000 0055 (BMD)
000 0130 (UMST500)
000 0132 (BMD)
multiple
000 0149 (UMST630)
nonlinear
000 0055 (BMD)
000 0132 (BMD)

partial
000 0149 (UMST630)
score
000 0073 (SCORMACH)

reliability
000 0007 (ATTANAL)
000 0073 (SCORMACH)
000 0111 (QUICKSCORE/ ITEMSTEP/ITEMRS)
test
000 0009 (GUTTSCL)
000 0035 (TSSA)
000 0095 (UOM 32)
000 0111 (QUICKSCORE/ ITEMSTEP/ITEMRS)

remote access
000 0101
000 0102 (RJE)
000 0103
000 0104
000 0105 (OLS)

Remote Job Entry
000 0102 (RJE)

replacements
000 0012 (BMD29)

reproducability coefficient
000 0007 (ATTANAL)
000 0009 (GUTTSCL)

residuals matrix
000 0007 (ATTANAL)
000 0012 (BMD29)
000 0078 (MOUFLON)
000 0087 (MINRES)

continued

- response(s)
  - frequency
    - 000 0009 (GUTTSCL)
    - 000 0055 (BMD)
    - 000 0132 (BMD)
  - patterns—Green, Loevinger, Sagi
    - 000 0009 (GUTTSCL)
  - scores
    - 000 0002 (FACTOR)
  - test
    - 000 0004 (SUZYQ)
    - 000 0035 (TSSA)
    - 000 0073 (SCORMACH)
    - 000 0095 (UOM 32)
- retrieval, information
  - 000 0024 (INFOL)
  - 000 0034 (TRIAL)
  - 000 0059 (BIRS)
  - 000 0099
- rho coefficient
  - 000 0031 (RKSTAT)
  - 000 0051 (m) (SPRHO)
  - 000 0096 (UOM 4)
- Richardson: see Kuder
- RJE system
  - 000 0102 (RJE)
- rotation, varimax
  - 000 0002 (FACTOR)
  - 000 0004 (SUZYQ)
  - 000 0027 (MESA1)
  - 000 0035 (TSSA)
  - 000 0051 (j) (VARMX)
  - 000 0066 (UMLFA)
- Runge—Kutta
  - 000 0054 (BEEFM)
  - 000 0082 (NODE)
  - 000 0083 (DNODE)

S

- Sagi response pattern
  - 000 0009 (GUTTSCL)
- sampling
  - 000 0071 (SPURT)
- Savage—Deutsch transaction flow
  - 000 0008 (INDIFF)
- scalability
  - 000 0009 (GUTTSCL)
- scaling
  - 000 0039 (SYMAP)
  - 000 0047 (SYMAP)
  - 000 0048 (QDGS)
  - 000 0052
  - 000 0069 (PLOT)
  - 000 0070 (MSA—I)
  - 000 0132 (BMD)
  - multidimensional
    - 000 0061 (MATDEC)
    - 000 0068 (KRUSKAL)
    - 000 0075 (KRUSCAL)
- scalogram
  - 000 0009 (GUTTSCL)
  - 000 0070 (MSA—I)
  - 000 0071 (SPURT)
- scattering
  - 000 0064 (ORFLS-PX)
- scheduling
  - 000 0093 (MPS/360)
  - 000 0133 (GPSS)
  - 000 0134 (CSMP)
  - network
    - 000 0029 (PERTC)
    - 000 0057 (PERT)
    - 000 0071 (SPURT)
  - student
    - 000 0090 (SCHEDULE)
- Scheffe's test
  - 000 0051 (p) (ANOVES)
- Schuessler
  - 000 0009 (GUTTSCL)
- scores, scoring
  - 000 0002 (FACTOR)
  - 000 0019 (CANON)
  - 000 0035 (TSSA)
  - 000 0073 (SCORMACH)
  - 000 0095 (UOM 32)
  - 000 0097 (UOM 87)
  - 000 0141 (UMST550)

continued

- scores, discriminant
  - 000 0010 (EIDISC)
  - psychometric
    - 000 0007 (ATTANAL)
    - 000 0009 (GUTTSCL)
  - response
    - 000 0004 (SUZYQ)
    - 000 0035 (TSSA)
    - 000 0073 (SCORMACH)
    - 000 0095 (UOM 32)
  - T
    - 000 0004 (SUZYQ)
    - 000 0045 (QSASE)
    - 000 0051 (m) (SPRHO)
    - 000 0095 (UOM 32)
  - test
    - 000 0004 (SUZYQ)
    - 000 0035 (TSSA)
    - 000 0095 (UOM 32)
    - 000 0097 (UOM 87)
    - 000 0111 (QUICKSCORE/ ITEMSTEP/ITEMRS)
  - U
    - 000 0001 (MANNWH)
    - 000 0087 (MINRES)
  - Z
    - 000 0001 (MANNWH)
    - 000 0004 (SUZYQ)
    - 000 0051 (1) (KETAU)
    - 000 0051 (n) (MNWHT)
    - 000 0111 (c) (ITEMRS)
    - 000 0139 (UMST540)
    - 000 0145 (UMST590)
- sensitivities
  - 000 0043 (PALS)
- sequence
  - 000 0023 (IDSEQN)
  - 000 0032 (SEQCHK)
- serials
  - 000 0140 (UPDATE/HOLD/ PUBLSH/LANSUB)
- significance
  - 000 0001 (MANNWH)
  - 000 0007 (ATTANAL)
  - 000 0030 (PPRANK)
- significance *(cont.)*
  - 000 0042 (CANNON)
  - 000 0051 (e) (SIGPP)
  - 000 0062 (ZFE-03, ZFE-04)
  - 000 0087 (MINRES)
- simplex
  - 000 0025 (LINPROG)
  - 000 0056 (LP1107)
  - 000 0077 (ZORILLA)
- simulation
  - 000 0037 (SAMOS)
  - 000 0071 (SPURT)
  - 000 0133 (GPSS)
  - 000 0134 (CSMP)
- single–factoredness
  - 000 0007 (ATTANAL)
- skewness
  - 000 0035 (TSSA)
  - 000 0051 (b) (STSUM)
  - 000 0117 (NORMSURV)
  - Pearsonian
    - 000 0008 (INDIFF)
- smallest space analysis
  - 000 0113 (SSA-1)
- sociology: see Behavioral Sciences
- Somers D
  - 000 0028 (NUCROS)
  - 000 0116 (NUCROS)
  - 000 0145 (UMST590)
- sort(ing)
  - 000 0006 (INDEXER)
  - 000 0034 (TRIAL)
  - 000 0038 (CONTEXT, CONCORD)
  - 000 0053 (BEEFDP)
  - 000 0084 (ABSRT)
  - 000 0085 (DBSRT)
- Spearman
  - 000 0007 (ATTANAL)
  - 000 0030 (PPRANK)
  - 000 0031 (RKSTAT)
  - 000 0051 (m) (SPRHO)
  - 000 0096 (UOM 4)
  - 000 0139 (UMST540)

continued

standard
  deviation
    000 0002 (FACTOR)
    000 0007 (ATTANAL)
    000 0008 (INDIFF)
    000 0010 (EIDISC)
    000 0014 (NUCORR)
    000 0018 (BISR)
    000 0026 (MANOVA)
    000 0027 (MESA1)
    000 0035 (TSSA)
    000 0045 (QSASE)
    000 0051 (b) (STSUM)
    000 0051 (d) (PPMCR)
    000 0051 (h) (UPREG/DNREG)
    000 0071 (SPURT)
    000 0073 (SCORMACH)
    000 0089 (ACT version 1.00)
    000 0095 (UOM 32)
    000 0098 (UOM 5)
    000 0106 (SUMSCRDS)
    000 0107 (REGAN1)
    000 0108 (SABCA)
    000 0109 (NUMFREQ)
    000 0110 (MISREGN)
    000 0111 (a) (QUICKSCORE)
    000 0117 (NORMSURV)
    000 0129 (STUDENTT)
    000 0130 (UMST500)
    000 0141 (UMST550)
    000 0142 (UMST560)
    000 0146 (UMST600)
    000 0149 (UMST630)

  errors
    000 0012 (BMD29)
    000 0027 (MESA1)
    000 0033 (TIMEX)
    000 0035 (TSSA)
    000 0036 (Z3SLS)
    000 0045 (QSASE)
    000 0051 (b) (STSUM)
    000 0051 (c) (TTEST)
    000 0055 (BMD)
    000 0071 (SPURT)
    000 0078 (MOUFLON)
    000 0107 (REGAN1)
    000 0109 (NUMFREQ)

standard errors *(cont.)*
    000 0110 (MISREGN)
    000 0129 (STUDENTT)
    000 0130 (UMST500)
    000 0141 (UMST550)
    000 0142 (UMST560)
    000 0149 (UMST630)

statistic
  discriminant
    000 0067 (DISCRIM2)
  distance
    000 0010 (EIDISC)
  F
    000 0129 (STUDENTT)
  Hocking–Leslie
    000 0078 (MOUFLON)
  probability
    000 0029 (PERTC)
  rank-order
    000 0139 (UMST540)
  T
    000 0130 (UMST500)
  U
    000 0001 (MANNWH)
    000 0031 (RKSTAT)
    000 0051 (n) (MNWHT)
    000 0087 (MINRES)

Statistics & Measurements
    000 0001 (MANNWH)
    000 0002 (FACTOR)
    000 0003 (STDNTF2)
    000 0004 (SUZYQ)
    000 0005 (MULCVR)
    000 0007 (ATTANAL)
    000 0009 (GUTTSCL)
    000 0010 (EIDISC)
    000 0011 (PLOTYY)
    000 0012 (BMD29)
    000 0013 (STDNTF1)
    000 0014 (NUCORR)
    000 0016 (AID)
    000 0017 (B34T)
    000 0018 (BISR)
    000 0019 (CANON)
    000 0020 (CHISQR)
    000 0021 (DISCRIM)
    000 0022 (EIGSYS)

continued

Statistics & Measurements *(cont.)*

000 0026 (MANOVA)
000 0027 (MESA1)
000 0028 (NUCROS)
000 0029 (PERTC)
000 0030 (PPRANK)
000 0031 (RKSTAT)
000 0033 (TIMEX)
000 0035 (TSSA)
000 0036 (Z3SLS)
000 0042 (CANON)
000 0044 (PROFILE)
000 0045 (QSASE)
000 0046 (INOUT)
000 0051 (a) (FAWCS)
000 0051 (b) (STSUM)
000 0051 (c) (TTEST)
000 0051 (d) (PPMCR)
000 0051 (e) (SIGPP)
000 0051 (f) (PARCOR)
000 0051 (g) (CANON)
000 0051 (h) (UPREG/DNREG)
000 0051 (i) (FANAL)
000 0051 (j) (VARMX)
000 0051 (k) (PHICO)
000 0051 (l) (KETAU)
000 0051 (m) (SPRHO)
000 0051 (n) (MNWHT)
000 0051 (o) (KRWAL)
000 0051 (p) (ANOVUM)
000 0051 (q) (ANOVES)
000 0051 (r) (AOVRM)
000 0051 (s) (BARTL)
000 0055 (BMD)
000 0060 (MATCHFS)
000 0061 (MATDEC)
000 0062 (ZFE-03, ZFE-04)
000 0063 (MANOVA)
000 0065 (OBLIMIN)
000 0066 (UMLFA)
000 0067 (DISCRIM2)
000 0068 (KRUSKAL)
000 0070 (MSA—I)
000 0071 (SPURT)
000 0073 (SCORMACH)
000 0074 (MSA)
000 0075 (KRUSCAL)
000 0076 (GAVIAL)

Statistics & Meas. *(cont.)*

000 0077 (ZORILLA)
000 0078 (MOUFLON)
000 0081 (FURNIVAL)
000 0087 (MINRES)
000 0089 (ACT version 1.00)
000 0094 (NYBMUL)
000 0095 (UOM 32)
000 0096 (UOM 4)
000 0097 (UOM 87)
000 0098 (UOM 5)
000 0106 (SUMSCRDS)
000 0107 (REGAN1)
000 0108 (SABCA)
000 0109 (NUMFREQ)
000 0110 (MISREGN)
000 0111 (QUICKSCORE/ITEMSTEP/ITEMRS)
000 0113 (SSA-1)
000 0114 (LSPOL)
000 0115 (STEPREGN)
000 0116 (NUCROS)
000 0117 (NORMSURV)
000 0128 (TARSIER)
000 0129 (STUDENTT)
000 0130 (UMST500)
000 0132 (BMD)
000 0133 (GPSS)
000 0137 (UMST520)
000 0138 (UMST530)
000 0139 (UMST540)
000 0141 (UMST550)
000 0142 (UMST560)
000 0143 (UMST570)
000 0144 (UMST580)
000 0145 (UMST590)
000 0146 (UMST600)
000 0147 (UMST620)
000 0148 (UMST610)
000 0149 (UMST630)

stepwise procedure

000 0010 (EIDISC)
000 0017 (B34T)
000 0051 (h) (UPREG/DNREG)
000 0055 (BMD)
000 0078 (MOUFLON)
000 0111 (a) (QUICKSCORE)
000 0115 (STEPREGN)

continued

stepwise procedure *(cont.)*
000 0132 (BMD)
000 0144 (UMST580)

stochastic
000 0071 (SPURT)

storage, information
000 0024 (INFOL)
000 0034 (TRIAL)
000 0059 (BIRS)

stress
000 0075 (KRUSCAL)

structure factors
000 0064 (ORFLS-PX)

Strutt: see Rayleigh

Student's
F ratio
000 0003 (STDNTF2)
T
000 0003 (STDNTF2)
000 0004 (SUZYQ)
000 0013 (STDNTF1)
000 0129 (STUDENTT)
000 0139 (UMST540)

synagraphic
000 0039 (SYMAP)
000 0047 (SYMAP)
000 0080 (SYMAP)

T

tau, Kendall
000 0028 (NUCROS)
000 0051 (1) (KETAU)
000 0116 (NUCROS)
000 0139 (UMST540)
000 0145 (UMST590)

teletypes
000 0101
000 0102 (RJE)
000 0103
000 0136

temperature factor coefficient
000 0064 (ORFLS-PX)

terminal
000 0101
000 0102 (RJE)
000 0103
000 0104
000 0105 (OLS)
000 0136

test
000 0004 (SUZYQ)
000 0016 (AID)
000 0019 (CANON)
000 0030 (PPRANK)
000 0035 (TSSA)
000 0042 (CANON)
000 0073 (SCORMACH)
analysis
000 0004 (SUZYQ)
000 0035 (TSSA)
000 0073 (SCORMACH)
000 0095 (UOM 32)
000 0111 (QUICKSCORE/ITEMSTEP/ITEMRS)
Bartlett's
000 0042 (CANON)
000 0051 (p) (ANOVUM)
000 0051 (q) (ANOVES)
000 0051 (s) (BARTL)
chi suqare
000 0020 (CHISQR)
evaluation
000 0041 (DOVACK)
F
000 0003 (STDNTF2)
000 0033 (TIMEX)
000 0045 (QSASE)
000 0078 (MOUFLON)
000 0149 (UMST630)
Hotelling's T
000 0005 (MULCVR)
Kendall
000 0028 (NUCROS)
000 0030 (PPRANK)
Mood
000 0033 (TIMEX)

continued

test, recognition
000 0041 (DOVACK)
reliability
000 0007 (ATTANAL)
000 0009 (GUTTSCL)
000 0035 (TSSA)
000 0073 (SCORMACH)
000 0095 (UOM 32)
000 0111 (QUICKSCORE/ ITEMSTEP/ITEMRS)
Scheffe's
000 0051 (p) (ANOVUM)
score
000 0004 (SUZYQ)
000 0035 (TSSA)
000 0095 (UOM 32)
000 0097 (UOM 4)
000 0111 (QUICKSCORE/ ITEMSTEP/ITEMRS)
significance
000 0062 (ZFE-03. ZFE-04)
significance. Lawley
000 0007 (ATTANAL)
T
000 0001 (MANNWH)
000 0003 (STDNTF2)
000 0005 (MULCVR)
000 0013 (STDNTF1)
000 0033 (TIMEX)
000 0051 (b) (STSUM)
000 0051 (m) (SPRHO)
000 0096 (UOM 4)
000 0098 (UOM 5)
000 0129 (STUDENTT)
000 0130 (UMST500)
000 0149 (UMST630)
U
000 0001 (MANNWH)
000 0031 (RKSTAT)
000 0051 (n) (MNWHT)
W
000 0030 (PPRANK)
Z
000 0001 (MANNWH)

tetrachoric
000 0004 (SUZYQ)
000 0035 (TSSA)

T
Hotelling's
000 0005 (MULCVR)
score
000 0004 (SUZYQ)
000 0033 (TIMEX)
000 0045 (QSASE)
000 0051 (m) (SPRHO)
000 0095 (UOM 32)
Student's
000 0003 (STDNTF2)
000 0004 (SUZYQ)
000 0013 (STDNTF1)
000 0129 (STUDENTT)
000 0139 (UMST540)
test
000 0001 (MANNWH)
000 0013 (STDNTF1)
000 0035 (TSSA)
000 0051 (b) (STSUM)
000 0096 (UOM 4)
000 0098 (UOM 5)
000 0129 (STUDENTT)
000 0149 (UMST630)

text processing
000 0034 (TRIAL)
000 0038 (CONTEXT, CONCORD)
000 0058 (I/II)
000 0079 (TEXT360)

time series
000 0033 (TIMEX)
000 0055 (BMD)
000 0132 (BMD)

time-sharing
000 0101
000 0102 (RJE)
000 0103
000 0104
000 0105 (OLS)
000 0136

topography
000 0039 (SYMAP)
000 0047 (SYMAP)
000 0080 (SYMAP)

continued

- traffic
  - 000 0040 (TRAN/PLAN)
- transaction
  - 000 0133 (GPSS)
  - Savage–Deutsch
    - 000 0008 (INDIFF)
- transformation
  - 000 0011 (PLOTYY)
  - 000 0012 (BMD29)
  - 000 0026 (MANOVA)
  - 000 0048 (QDGS)
  - 000 0063 (MANOVA)
  - 000 0094 (NYBMUL)
  - 000 0117 (NORMSURV)
  - 000 0141 (UMST550)
  - 000 0143 (UMST570)
  - 000 0144 (UMST580)
  - 000 0149 (UMST630)
  - QR
    - 000 0043 (PALS)
    - 000 0081 (FURNIVAL)
    - 000 0086 (CGELG)
- transgeneration
  - 000 0011 (PLOTYY)
  - 000 0012 (BMD29)
  - 000 0055 (BMD)
  - 000 0129 (STUDENTT)
  - 000 0132 (BMD)
  - 000 0141 (UMST550)
  - 000 0143 (UMST570)
  - 000 0145 (UMST590)
  - 000 0149 (UMST630)
- transportation
  - planning
    - 000 0040 (TRAN/PLAN)
  - problem
    - 000 0135 (TRANSPRT)
- treatments
  - 000 0062 (ZFE-03, ZFE-04)
- trips
  - 000 0040 (TRAN/PLAN)

U

- unidimensionality
  - 000 0007 (ATTANAL)
- U statistic
  - 000 0001 (MANNWH)
  - 000 0031 (RKSTAT)
  - 000 0051 (n) (MNWHT)
  - 000 0087 (MINRES)
- utility program
  - 000 0050 (a) (FMACUT)

V

- variance
  - 000 0033 (TIMEX)
  - 000 0066 (UMLFA)
  - 000 0067 (DISCRIM2)
  - 000 0098 (UOM 5)
  - 000 0108 (SABCA)
  - 000 0111 (QUICKSCORE/ ITEMSTEP/ITEMRS)
  - 000 0129 (STUDENTT)
  - 000 0138 (UMST530)
  - 000 0141 (UMST550)
  - 000 0142 (UMST560)
  - 000 0146 (UMST600)
  - 000 0148 (UMST610)
  - 000 0149 (UMST630)
  - analysis
    - 000 0004 (SUZYQ)
    - 000 0005 (MULCVR)
    - 000 0007 (ATTANAL)
    - 000 0010 (EIDISC)
    - 000 0016 (AID)
    - 000 0026 (MANOVA)
    - 000 0027 (MESA1)
    - 000 0051 (f) (PARCOR)
    - 000 0051 (i) (FANAL)
    - 000 0051 (p) (ANOVUM)
    - 000 0051 (q) (ANOVES)
    - 000 0051 (r) (AOVRM)
    - 000 0055 (BMD)
    - 000 0063 (MANOVA)

continued

variance analysis *(cont.)*
000 0076 (GAVIAL)
000 0087 (MINRES)
000 0094 (NYBMUL)
000 0108 (SABCA)
000 0132 (BMD)
000 0143 (UMST570)
000 0148 (UMST610)

varimax
  matrix
    000 0141 (UMST550)
  rotation
    000 0002 (FACTOR)
    000 0004 (SUZYQ)
    000 0027 (MESA1)
    000 0035 (TSSA)
    000 0051 (j) (VARMX)
    000 0066 (UMLFA)

vocabulary
000 0041 (DOVACK)

Von Neumann
000 0012 (BMD29)
000 0055 (BMD)
000 0132 (BMD)

V, Rao's
000 0010 (EIDISC)

W

Walker-Lev test
000 0033 (TIMEX)

Wallis: see Kruskal

Watson, Durbin–
000 0012 (BMD29)
000 0130 (UMST500)
000 0149 (UMST630)

weight(ing)
000 0004 (SUZYQ)
000 0019 (CANON)
000 0039 (SYMAP)
000 0042 (CANON)
000 0047 (SYMAP)
000 0080 (SYMAP)
000 0097 (UOM 87)
000 0142 (UMST560)

weight(ing) *(cont.)*
000 0143 (UMST570)
000 0144 (UMST580)

Whitney: see Mann

Wilk lambda
000 0010 (EIDISC)

word list(s)
000 0041 (DOVACK)
000 0038 (CONTEXT, CONCORD)
000 0049 (KGIC)

W
  index
    000 0111 (b) (ITEMSTEP)
    000 0111 (c) (ITEMRS)
  Kendall coefficient
    000 0030 (PPRANK)
    000 0031 (RKSTAT)
    000 0139 (UMST540)
  test
    000 0030 (PPRANK)

X

Y

Yates correction factor
000 0051 (a) (FAWCS)
000 0137 (UMST520)
000 0145 (UMST590)
000 0147 (UMST620)

Youden rectangles
000 0062 (ZFE-03, ZFE-04)

Z

zoology: see Life Sciences

Z score
000 0001 (MANNWH)
000 0004 (SUZYQ)
000 0051 (l) (KETAU)
000 0051 (n) (MNWHT)
000 0111 (c) (ITEMRS)
000 0139 (UMST540)
000 0145 (UMST590)